Environment Canada
Environnement Canada

PULP AND PAPER TECHNICAL GUIDANCE FOR AQUATIC ENVIRONMENTAL EFFECTS MONITORING

EEM/1998/1
APRIL 1998

ENVIRONMENT CANADA

AUSSI DISPONIBLE EN FRANÇAIS

Canada

Publications relating to Environmental Effects Monitoring (EEM):

Aquatic Environmental Effects Monitoring Requirements (revised EPS 1/RM/18). Environment Canada. 1997. EEM/1997/1

Pulp and Paper Aquatic Environmental Effects Monitoring Requirements. Environment Canada. 1997. Annex 1 to EEM/1997/1

Tracer Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/2 is also published as: Ali, N et al, 1997. Exposed or not exposed: the effectiveness of using tracers to verify exposure of fish to pulp and paper mill effluent. Canadian Technical Report of Fisheries and Aquatic Sciences 2162.

Dioxin Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/3

Usability Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/4

Toxicology Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/5

Fish Survey Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/6

Benthic Community Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/7

Lowell, R.B. Discussion paper on critical effect size guidelines for EEM using benthic invertebrate communities. Report to the Environmental Effects Monitoring Program. EEM/1997/8

Pulp and Paper Technical Guidance for Aquatic Environmental Effects Monitoring. 1998. EEM/1998/1

Copies are available at:

National EEM Office
Environmental Quality Branch
Environmental Conservation Service
Environment Canada
Ottawa, Ontario, Canada K1A 0H3
Tel.: (819) 997-1535
Fax: (819) 994-1691
Email: eem@ec.gc.ca
Website: http://www.ec.gc.ca/eem/english/default.htm

Table of Contents

1 **Introduction** 1-1

2 **EEM Pre-Design** 2-1
2.1 Introduction 2-1
2.2 Description of Study Area 2-3
2.2.1 Delineation of Zone of Effluent Mixing 2-3
2.2.1.1 Plume Delineation Studies 2-4
2.2.2 Using Tracers 2-5
2.2.2.1 Type of Tracers 2-5
2.2.2.2 Mixing 2-5
2.2.2.3 Degradation of Tracers 2-6
2.2.2.4 Injection System 2-6
2.2.2.5 Measurements in the Field 2-7
2.2.3 Using Predictive Models 2-8
2.2.4 Environment-Specific Considerations 2-9
2.2.4.1 Rivers 2-9
2.2.4.2 Lakes 2-10
2.2.4.3 Estuaries 2-11
2.2.4.4 Fjords 2-12
2.2.4.5 Oceans 2-12
2.3 Habitat Inventory and Classification 2-13
2.3.1 Environment-Specific Considerations 2-15
2.3.2 Habitat Classification 2-15
2.3.3 Delineation of Depositional Zones 2-16
2.4 Resource Inventory 2-19
2.5 Detailed Site Map 2-20
2.6 Historical Receiving Environment Data 2-20
2.7 Effluent Quality and Characterization 2-21
2.8 Mill History and Operations 2-22
2.9 References 2-24

3 **Sampling Design** 3-1
3.1 Description of Study Areas 3-1
3.2 Selection of Reference Areas in Marine and Estuarine Environments 3-2
3.3 Selection of Reference Areas in Freshwater Environments 3-3
3.3.1 The Control-Impact (C-I) Approach in Reference Areas Selection 3-3
3.3.2 Classification Scheme for Reference Area Selection 3-4
3.3.2.1 Ecoregions 3-4
3.3.2.2 Drainage Basin and Geographic Scale 3-7
3.3.2.3 Land-use and Vegetative Buffer Strips 3-7
3.3.3 Reference Area Selection in Rivers 3-8

3.3.3.1 Valley Segment and Stream Order ... 3-8
3.3.3.2 Stream Reach ... 3-8
3.3.3.3 Channel Unit ... 3-10
3.3.4 Reference Site Selection in Lakes ... 3-11
3.3.4.1 Origins, Location and Hydrological Linkages ... 3-12
3.3.4.2 Morphometry ... 3-12
3.3.4.3 Trophic State ... 3-12
3.3.4.4 Zone ... 3-13
3.4 Reporting of Field Station Positions ... 3-13
3.5 Statistical Considerations ... 3-14
3.5.1 Hypotheses Testing ... 3-14
3.5.2 Components of Variation ... 3-15
3.5.3 Power Analyses and Critical Effect Size Considerations ... 3-16
3.6 References ... 3-18

4 General Quality Assurance and Quality Control (QA/QC) for Conducting EEM ... 4-1
4.1 Quality Management ... 4-5
4.1.1 Standard Operating Procedures ... 4-5
4.1.2 Data Quality Objectives ... 4-6
4.2 Quality Assurance Functions ... 4-6
4.3 Field Operations ... 4-6
4.4 References ... 4-8

5 Effects on Fish and Fisheries Resources: Fish Survey, Tissue Analyses and Tainting Evaluation ... 5-1
5.1 Introduction to the Fish Survey ... 5-1
5.1.1 Fish Survey ... 5-1
5.1.1.1 Approach ... 5-1
5.1.1.2 Previous Cycles ... 5-3
5.1.2 Selection of Sentinel Monitoring Species ... 5-9
5.1.3 Sampling Design ... 5-12
5.1.3.1 Minimum Design ... 5-12
5.1.3.2 Timing of Collection ... 5-15
5.1.3.3 Selection of Sampling Areas ... 5-17
5.1.3.4 Fish Sampling Methods ... 5-20
5.1.3.5 Sampling Effort ... 5-23
5.1.4 Alternatives to the Fish Survey ... 5-27
5.1.5 Data Analysis ... 5-27
5.1.5.1 Variables to be Measured ... 5-27
5.1.5.2 Data Handling ... 5-31
5.1.5.3 Data Summary and Screening ... 5-31
5.1.5.4 Assessing Normality and Equality of Variance ... 5-31
5.1.5.5 Data Transformation ... 5-32

5.1.6 Fish Parameters 5-32
5.1.6.1 Principal Parameters and Estimation of Variability 5-32
5.1.6.2 Use of Indices 5-36
5.1.7 Variability and Sample Size Considerations 5-36
5.1.8 Reporting 5-38
5.2 Tissue Analyses: Chlorinated Dioxins and Furans Congeners 5-44
5.2.1 Tissue Sample Collection and Preparation 5-46
5.2.2 QA/QC Criteria for Clorinated Dioxins and Furans Analyses 5-46
5.2.3 Lipid Analysis 5-49
5.3 Tainting Evaluation 5-50
5.3.1 Introduction 5-50
5.3.2 Sample Collection and Sample Handling Methods 5-53
5.3.2.1 Sample Collection 5-53
5.3.2.2 Sample Handling 5-61
5.3.3 Sample Preparation and Sensory Evaluation Methods 5-62
5.3.3.1 Background 5-62
5.3.3.2 Sample Preparation 5-64
5.3.3.3 Selection, Training and Monitoring of a Sensory Evaluation Panel 5-66
5.3.4 Experimental Design and Statistical Analysis 5-75
5.3.5 Reporting Requirements 5-77
5.4 References 5-79

6 Effects on Fish Habitat: Invertebrate Community Survey 6-1
6.1 Study Design 6-2
6.1.1 Standard Nomenclature 6-2
6.1.2 Sampling Locations 6-3
6.1.3 Pre-operational and EEM Cycle 1 Invertebrate Community Surveys 6-6
6.1.4 EEM Cycle 2 Invertebrate Community Surveys 6-7
6.1.5 Sampling Design 6-15
6.1.5.1 Selection of Reference Areas 6-15
6.1.5.2 Sampling Periods 6-19
6.1.5.3 Selection of Dominant Habitat Type 6-19
6.1.5.4 Sampling Program Designs 6-23
6.1.5.5 Sampling Effort 6-32
6.2 Benthic Invertebrate Sampling Methods 6-38
6.2.1 Field Methods 6-38
6.2.1.1 Field Sieves and Mesh Sizes 6-38
6.2.1.2 Sampling Equipment 6-40
6.2.1.3 Specimen Preservation 6-46
6.2.1.4 QA/QC for Benthic Invertebrate Field Operations 6-46
6.2.1.5 Supporting Environmental Variables 6-47
6.2.2 Laboratory Methods 6-49
6.2.2.1 Sample Sorting 6-49

6.2.2.2 Laboratory Subsampling 6-50
6.2.2.3 Taxonomy 6-51
6.2.2.4 Establishment of Reference Collections 6-52
6.2.2.5 Invertebrate Biomass 6-55
6.2.2.6 QA/QC for Benthic Invertebrate Laboratory Operations 6-55
6.3 Data Analysis 6-57
6.3.1 QA/QC for Data Input and Verification 6-57
6.3.2 Data Handling 6-58
6.3.3 Statistical Analyses 6-61
6.3.4 Reporting Requirements 6-64
6.4 References 6-66

7 Measurement of Supporting Environmental Variables 7-1
7.1 Introduction 7-1
7.2 Receiving Water Monitoring 7-1
7.2.1 Sample Collection 7-1
7.2.2 Sample Storage and Shipping 7-5
7.2.3 Toxicity 7-5
7.3 Sediment 7-5
7.3.1 Freshwater Environments 7-5
7.3.2 Marine Environments 7-6
7.3.3 Critical Effects for Sediment Chemistry in Marine/Estuarine Environments 7-8
7.4 Quality Assurance in Analyses 7-11
7.4.1 Sample Handling and Tracking 7-11
7.4.2 Good Laboratory Practices 7-13
7.4.3 GC/MS (Gas Chromatrography/Mass Spectroscopy) Tuning and Performance Checks 7-13
7.4.4 Organic Analytes Identification 7-14
7.4.5 Instrument Response Checks 7-14
7.4.6 Instrument Calibration Requirements 7-14
7.4.7 Analytical Documentation 7-20
7.5 Analytical Quality Control 7-21
7.5.1 Calibration Control 7-21
7.5.2 Precision 7-23
7.5.3 Accuracy and Bias 7-24
7.5.4 Recovery 7-25
7.5.5 Detection Limits 7-27
7.6 Data Reporting Conventions 7-27
7.7 Method Principles - Physical and Chemical 7-28
7.7.1 Resin and Fatty Acids in Effluent 7-29
7.7.2 Chlorinated Dioxins and Furans in Sediments 7-29
7.7.3 Chlorinated Phenols, Guaiacols, and Catechols 7-29
7.8 Chemical Tracers 7-29

7.8.1 Applicability of Tracers 7-30
7.8.2 Selection of Tracers 7-30
7.8.3 Sample Collection and Analysis 7-34
7.8.4 Tracers in Benthos 7-36
7.9 References 7-37

8 Process Effluent Sublethal Toxicological Testing 8-1
8.1 Background 8-1
8.2 Tabulation of Endpoints and Validation of Test Results 8-1
8.2.1 Validation of Test Results 8-2
8.2.2 Tabulation of Endpoints 8-5
8.3 Data Interpretation in Relation to the Toxicity Objectives 8-6
8.3.1 Changes in Effluent Quality 8-6
8.3.2 Understanding Multiple Discharge Situations 8-6
8.3.3 Contribution to the Weight of Evidence Approach 8-6
8.3.4 Considerations for Integration of Toxicity Test Results 8-7
8.4 References 8-8

9 Recommended Outline for Interpretative Reports 9-1
9.1 Introduction 9-1
9.2 Information Requirements 9-1
9.2.1 Site Description and Mill Update 9-1
9.2.2 EEM Study 9-2
9.3 References 9-4

10 Information Management 10-1

Glossary G-1

List of Tables

Table 2.1 Pre-design information requirements for EEM 2-2
Table 2.2 Reporting requirements for resources and habitat inventory of the study area.. 2-14
Table 2.3 Substrate subclasses and dominance types (after *Cowardin et al.* 1979) 2-18
Table 2.4 Salinity modifiers (after *Cowardin et al.* 1979) 2-18
Table 2.5 Water regime modifiers (after *Cowardin et al.* 1979) 2-18
Table 2.6 Example of resource inventory information 2-20
Table 2.7 Example of annual summary of mill effluent characteristics 2-22
Table 2.8 Example of pulp and paper mill operations checklist 2-23
Table 4.1 Example of quality assurance checklist for aquatic biology studies 4-2
Table 4.2 Example of quality assurance checklist for aquatic toxicology studies 4-3
Table 4.3 Example of quality assurance checklist for physical and chemical analyses 4-4
Table 5.1 Fish survey measurements 5-2
Table 5.2 Sentinel species characteristics and criteria for success 5-6
Table 5.3 Were fish exposed? 5-6
Table 5.4 Guidance for re-design 5-7
Table 5.5 Are there sufficient recent data for a conclusion? 5-7
Table 5.6 Other concerns which could affect a survey 5-8
Table 5.7 Power Analysis 5-8
Table 5.8 Possible alternatives for the fish survey 5-8
Table 5.9 Anatomical structures commonly used for age determination 5-11
Table 5.10 Sample checklist for physical examination of fishes 5-29
Table 5.11 Definition of fish parameters used in the Fish Survey (from *Gibbons et al.* 1992) 5-34
Table 5.12 Suggested reporting format for the parameters (A) and the resulting regressions (B) required for the fish survey analysis 5-43
Table 5.13 Performance criteria for analysis of chlorinated dioxins and furans 5-48
Table 5.14 Data from degree of difference test 5-77
Table 5.15 Analysis of variance for Table 5.14 data 5-77
Table 6.1 In the previous cycle, did the benthic community differ between exposed and reference areas or was there a gradient response? 6-10
Table 6.2 Was the sampling design adequate to suggest that the effects were linked to mill effluent? Are confounding factors or sampling methods potentially influencing the results? 6-11
Table 6.3 Was the statistical power adequate to determine that there was no mill effect? 6-12
Table 6.4 Has relative effluent loading significantly changed since the last cycle? 6-13
Table 6.5 Was the sampling design adequate to determine the magnitude and spatial extent of the effect? 6-13
Table 6.6 Are approved alternatives available for the invertebrate survey? 6-14
Table 6.7 Are there sufficient recent data for a conclusion? 6-14
Table 6.8 Benthic sampling program designs 6-27
Table 6.9a Sample sizes required to detect difference of 1 SD 6-35

Table 6.9b Sample sizes required to detect defference of 2 SD 6-35
Table 6.10 Summary criteria for recommended samplers 6-44
Table 6.11 Supporting variables for freshwater (F) and marine/estuarine (M) habitats to be measured concurrently with benthic faunal surveys 6-48
Table 6.12 Recommended level of taxonomic precision for benthic macroinvertebrates ... 6-53
Table 6.13 List of freshwater and marine taxonomic macroinvertebrate keys 6-54
Table 6.14 Key references and sources 6-65
Table 7.1a Supporting environmental variables in water for freshwater (F) and marine (M) habitats 7-2
Table 7.1b Supporting environmental variables in sediment for freshwater (F) and marine (M) habitats 7-3
Table 7.1c Other supporting environmental variables 7-3
Table 7.2 Handling recommendations for effluent and receiving water samples 7-4
Table 7.3 Characteristics of coring devices (after *Klemm et al.* 1990) 7-7
Table 7.4 TOC sediment criteria for determining "normal" sediment conditions for monitoring in proximity to fish farms in Washington State 7-9
Table 7.5 Summary of analytical method principles for effluents and receiving waters... 7-15
Table 7.6 Summary of analytical method principles for sediment 7-18
Table 7.7 Recommended quality control for inorganic analyses 7-22
Table 7.8 Recommended quality control for organic analyses 7-22
Table 8.1 Methods for toxicity tests with pulp and paper process effluents (Environment Canada, 1997b) 8-2
Table 8.2 Minimum reporting requirements for *Ceriodaphnia dubia* sublethal toxicity test 8-4
Table 8.3 Hypothetical example of potential zone of effect estimates (assuming of 1% zone of 1500m) 8-7

List of figures

Figure 2.1 Classification hierarchy of wetlands and deep water habitats showing systems, subsystems, and classes (after *Cowardin et al.* 1979) 2-17
Figure 3.1 Classification scheme for reference site selection 3-5
Figure 3.2 Ecoregions of Canada (after Environment Canada 1997)........ 3-6
Figure 3.3 Spatially nested hierarchy for stream systems (after *Cowardin et al.* 1994) 3-9
Figure 3.4 Dendrogram indicating how channel units can be classified with increasing levels of resolution (after *Hawkins et la.* 1993) 3-11
Figure 3.5 Study areas risk (α) and environmental risk (β) 3-15
Figure 5.1 Fish survey decision tree........ 5-5
Figure 5.2 Changes in upper and lower 95% confidence limits of example standard deviations of 5 and 10, in relation to sample size 5-16
Figure 5.3 Features of composite fish 5-28
Figure 5.4 Dioxins and furans decision tree........ 5-45
Figure 5.5 Tainting decision tree........ 5-52
Figure 5.6 Example of a sensory evaluation ballot used for DFC tests with fish exposed to pulp and paper mill effluent (form Beak International)........ 5-65
Figure 5.7 Graph for sequential analysis for selection panel trainees using triangle tests.. 5-70
Figure 5.8 An example of the sequential approach to panel selection using fish exposed to three concentrations of pulp and paper mill effluent (taken from Eurocan Pulp and Paper Effluent Monitoring Program, 1997) 5-72
Figure 6.1 Examples of AREA, STATION and SAMPLE spatial scales for a basic control/impact design (see section 6.1.5.3) 6-5
Figure 6.2 Benthic survey decision tree 6-9
Figure 6.3 Control/Impact design for freshwater rivers and streams 6-28
Figure 6.4 Multiple control/impact design........ 6-29
Figure 6.5a Linear gradient design for stream or river mills 6-30
Figure 6.5b Linear gradient design for lake or coastal mills situated in narrows bays or fjords 6-30
Figure 6.6 Radial gradient design for lake or coastal mills........ 6-31
Figure 6.7 Combination design (C/I with downstream reference area) for river mills. Note that the mills are homogeneous in terms of environmental variables....... 6-31
Figure 7.1 Example of a chain-of-custody record (modified from US EPA 1988b)........... 7-12
Figure 7.2 Fish tracer decision tree 7-32

1 Introduction

The environmental effects monitoring (EEM) program was developed to assess the adequacy of Pulp and Paper Effluent Regulations under the federal *Fisheries Act*. The program has been designed to achieve national uniformity in monitoring of effects, while taking into consideration site-specific factors. Program requirements are given in "Aquatic Environmental Effects Monitoring Requirements EEM/1997/1" (Environment Canada, 1997a) and "Annex 1 to EEM/1997/1: Pulp and Paper Aquatic Environmental Effects Monitoring Requirements." (Environment Canada, 1997b).

This document has been revised from the original 1992 version based on recommendations from a series of Expert Working Groups, and an overall review of Cycle 1 of the Pulp and Paper EEM program (Tracer Expert Working Group Report, 1997; Dioxin Expert Working Group Report, 1997; Usability Expert Working Group Report, 1997; Toxicology Expert Working Group Report, 1997; Fish Survey Expert Working Group Report, 1997; Benthic Community Expert Working Group Report, 1997). Please note that this Technical Guidance Document is published in a format and, is viewed as a document, which can and will be updated regularly as new information and research becomes available. To ensure that the most recent updates are included in this document contact Environment Canada's, National EEM Office (phone: 819-997-1535; fax 819-953-0461; e-mail: eem@ec.gc.ca; website: http://www.ec.gc.ca/eem).

The purpose of the Technical Guidance Document is to provide further details of how to carry out the monitoring requirements outlined in Annex 1. Throughout this document the terms "should" and "must" are used regarding design and QA/QC aspects. EEM interpretative reports must document that the technical requirements that are preceded by a "must" have been satisfied. Requirements preceded by "should" are not absolute requirements but meeting these requirements will support the overall technical quality of the data. In all cases, it is necessary to discuss design with the Regional Authorisation Officers (RAOs).

References

Benthic Community Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/7

Dioxin Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/3

Environment Canada. 1997a. Aquatic Environmental Effects Monitoring Requirements (revised EPS 1/RM/18). EEM/1997/1

Environment Canada. 1997b. Pulp and Paper Aquatic Environmental Effects Monitoring Requirements. Annex 1 to EEM/1997/1

Fish Survey Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/6

Toxicology Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/5

Tracer Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/2 is also published as: Ali, N et al, 1997. Exposed or not exposed: the effectiveness of using tracers to verify exposure of fish to pulp and paper mill effluent. Canadian Technical Report of Fisheries and Aquatic Sciences 2162.

Usability Expert Working Group Report. EEM Technical Management Committee. 1997. EEM/1997/4

2 EEM Pre-Design

2.1 Introduction

An EEM pre-design study is required only once, provided that the mill process, effluent loading, the discharge location, or the receiving environment conditions have not changed significantly. Pre-design information should be summarized and updated as part of each EEM design submission. The pre-design requirements are listed in Table 2.1.

Pre-design information is necessary to:

1) delineate the spatial extent of the study area, including the zone of effluent mixing and representative reference areas;
2) confirm that the samples collected are representative of these areas at the time of field sampling;
3) describe habitat type at a level of resolution sufficient to allow siting of invertebrate and fish sampling stations;
4) document any potentially confounding or influencing factors that must be considered in the study design and interpretation of results;
5) provide knowledge of the relative abundance of fish in the study area and determine the feasibility of selecting two sentinel fish species; and,
6) describe the quality and use of fisheries resources in the receiving environment to assist in the setting of statistical criteria in subsequent EEM cycles.

Table 2.1 Pre-design information requirements for EEM

Description of Study Area (Receiving Environment)[1]
Delineation of Zone of Effluent Mixing
Habitat Inventory and Classification
Resource Inventory
Historical Receiving Environment Data (including previous EEM findings)
Effluent Quality[2]
pH
Flow
Conductivity
Biochemical Oxygen Demand (BOD)
Total Suspended Solids (TSS)
96-h Rainbow Trout Lethality
48-h *Daphnia magna* Lethality
Chlorinated Dioxin and Furan Congeners[3]
Mill History and Operations[4]

---

[1] To be updated and summarized prior to each EEM cycle. Delineation of the zone of effluent mixing may be required if there was a change in the location of the effluent discharge or a significant change in the effluent loading.

[2] Requirements under the Pulp and Paper Effluent Regulations (1992) except where otherwise indicated.

[3] Requirements for Regulations Respecting the Release of Polychlorinated Dibenzo-*para*-Dioxins and Polychlorinated Dibenzofurans from Pulp and Paper Mills (*Canadian Environmental Protection Act*).

[4] Mills are required to report on mill history and operations data as specified in the Pulp and Paper Effluent Regulations, (1992) and Regulations Respecting the Release of Polychlorinated Dibenzo-*para*-Dioxins and Polychlorinated Dibenzofurans from Pulp and Paper Mills. Recommendations on additional data which will assist in interpretation of EEM results are given in Annex 1 (Table 20), Section 5 of this document.

2.2 Description of Study Area

2.2.1 Delineation of Zone of Effluent Mixing

Delineation of the zone of effluent mixing is necessary to ensure that:
1) sampling stations are located within the zone of effluent mixing; and
2) reference areas are well beyond any potential zone of impact.

In most receiving waters, the location and dilution gradients of the zone of effluent mixing will be dynamic and affected by changes in the receiving water environment and the effluent discharge. It is recommended that the zone of effluent mixing be delineated where possible and practical for conditions of minimum dilution, maximum extent, and long-term average conditions. Points within a minimum dilution region are affected by short-term exposure to the greatest concentrations of the effluent. Points outside the maximum extent regions are not affected by the discharge and may be suitable as reference areas. The long-term average condition defines any long-term effect of the effluent discharge.

General principles for delineating the zone of effluent mixing can be elaborated; variations for a particular receiving water will be discussed under Environment-Specific Considerations (Section 2.2.4). Most of the environment-specific considerations relate to the timing of the plume delineation studies relative to the characteristics of the receiving waters.

The zone of effluent mixing will most often be determined by a plume delineation study. The spatial extent of the zone is determined by a specified effluent concentration. Pulp and paper studies use the limit of 1% (100:1 dilution). Information on the use of plume delineation studies is described in section 2.2.1.1.

In some cases, it may not be possible or practical to delineate the zone of effluent mixing by the above approach. Examples of such situations include:

- Sites where the position of the effluent plume is highly transitory (*i.e.* the plume rapidly changes location, volume, or direction of flow)

- Sites where there is high dilution (*i.e.* where the effluent is diluted very rapidly near the outfall and the locating of sampling stations in the mixing zone would be very difficult). In this case, it might be wise, where the effluent seems to dilute rapidly, to radiate out from the outfall and keep looking for evidence of the plume beyond the very low concentrations. Likewise, when the dispersion of the plume is transitory, its maximal dispersion should be delineated, i.e. to include all areas of potential exposure.

In such cases, alternative methods of determining the exposure area may be needed. Alternative methods should be proposed to the Regional Authorization Officer (RAO) and accompanied by a detailed description or reference to a detailed description. In general, such an alternative would entail a spatial characterization of the receiving environment such as information on sediment or water quality, toxicity and mapping of benthic community structure.

2.2.1.1 Plume Delineation Studies

It is recommended that the range of the effluent mixing zone be defined for both high and low flow (tide) periods.

Plume delineation studies are done by measuring effluent concentrations usually of tracers (see Section 2.2.2). Each type of receiving environment will have different types of characteristics to be considered in the design of tests (see Section 2.2.4). Ideally, the plume should be tracked to background.

The limits of the initial discharge zone (or edge) can be identified in the field as the points at which the effluent plume and receiving water are moving at the same velocity.

The recommended method to verify effluent dilution predictions is to measure at the middle of the initial discharge zone by sampling at the depth of highest concentration. By collecting three or four samples, the initial dilution can be calculated as the maximum measured tracer concentration divided by the tracer concentration in the effluent prior to discharge.

In some special cases, it may not be possible to measure the plume to the specified limit of resolution of effluent concentration. The extent of the exposure area should be limited or expanded as appropriate, based on best professional judgement. For example, in a river with low flow, or in a receiving water where there are other discharges close to the mill discharges, these factors will act to complicate concentration measurements and confound the results.

2.2.2 Using Tracers

2.2.2.1 Type of Tracers

The ideal plume tracer has a near zero background level and decay rate (a conservative substance), can be readily measured in the field at low concentrations, and can be released at a constant loading rate. These conditions are desirable because the extent of the plume at the time of measurement should reflect near steady-state conditions during the period of measurement, and detection at low levels is needed to achieve the degree of resolution.

i. **Added Tracers:** The preferred tracer is a fluorescent dye. The recommended dye is Rhodamine WT. This dye has been shown not to be a mutagen (Douglas *et al.* 1983), has low potential for toxicity and adverse effects in the aquatic environment (Parker 1973), is safe when handled with care, generally available, can be readily measured in the field at concentrations in the low $\mu g \cdot L^{-1}$ range, can be considered conservative in most cases, and typically has a near zero background level. Other tracers, such as those below, where demonstrated as scientifically valid, could be used.

ii. **Constituents in discharge:** The effluents from most mills contain a variety of constituents that could potentially be used as tracers for delineating the zone of effluent mixing. Some of these constituents can be highly variable in terms of their mass loading to the receiving water and are not easily measured in the field. Effluent colour, sodium, chloride, tannin-lignins, conductivity and chloroform have been used.

2.2.2.2 Mixing

To assure complete mixing before the effluent reaches the receiving water, the injection location should be well upstream from the effluent outfall. If the dye is injected into an existing outfall, the turbulence in the outfall downstream from the injection point is sufficient to ensure complete mixing of dye throughout the effluent. The dilution of dye in the outfall is such that any change in effluent density caused by the addition of dye is insignificant. The adjustment of dye density is only necessary when there is no existing outfall and dye is released from a boat (e.g. at a proposed outfall location). Small differences in temperature (2°C to 3°C) between the effluent and the dye can cause a significant density difference. If the tracer is not the same density as the effluent, it may separate from the effluent in the receiving water. The density of the dye Rhodamine WT should be adjusted with methyl hydrate or a similar solvent (dilution liquid) that has a specific gravity of less than 1.

2.2.2.3 Degradation of Tracers

Fluorescent tracers such as Rhodamine WT can be affected by oxidizing agents (e.g. bleaches, sulphides, sunlight, salt water, and microorganisms). Preliminary tests of dye-effluent interaction to describe the stability of the tracer and to determine the loss coefficient that should be used with the tracer are recommended.

The degradative effects of sunlight and microorganisms on the tracer can be determined by measuring the tracer concentration dynamics in test samples exposed to sunlight and by maintaining the effluent tracer mixture and receiving water at appropriate dilutions during the tracer studies. If the degradation of the dye is too large to be useful, another more conservative tracer may be required. Sulphur hexafluoride has been used and can be detected in very small concentrations using gas chromatography.

2.2.2.4 Injection System

A continuous flow constant rate injection system is preferred to simulate the operation of an outfall with continuous discharge loading. This type of injection system makes field measurements more reliable. The use of a constant rate positive displacement pump (e.g. Masterflex L/S digital pump) is recommended. A constant head device is also a good method for injecting dye. The pumping rate should be checked at least at the beginning and the end of the injection by direct measurement. Plume measurements can be repeated and mass balance checks can be carried out at cross plume transects in the plume. For a tracer with zero background levels (e.g. a dye), the injection rate can be calculated as follows:

$$q_{inj} = q_{eff} \times C_x \times 100 / [\%_{eff} \times C_o]$$

where:	q_{inj}	is the tracer injection rate
	q_{eff}	is the effluent flow rate
	C_x	is the tracer detection limit concentration
	$\%_{eff}$	is the limit of the plume in % effluent concentration
	C_o	is the concentration of the tracer in the injection mixture

The concentration of the tracer should be considered in the computation of the injection rate. The concentration of the dye in the stock supplied by the manufacturer is defined. Any dilution of this stock should be considered in the computation of the injection rate to determine the concentration of the dye in the mixture injected.

Using the equation above, the constant rate injection of the tracer is determined by measuring the effluent discharge rate and then determining, as appropriate, the tracer injection rate based on the detection limits of the tracer measuring technique. For example, if the detection limit of the tracer measuring method is 1 $\mu g \cdot L^{-1}$, the limit of resolution of the plume is 1% effluent concentration, the effluent flow is 1000 L·s and the tracer concentration in the injection mixture is 1×10^9 $\mu g \cdot L^{-1}$, then using the above equation the required pumping rate is 0.0001 L·s.

2.2.2.5 Measurements in the Field

The following section assumes the use of Rhodamine WT as the tracer. Use of an alternative tracer will require appropriate modifications to the recommended procedure. Details on the use of Rhodamine WT are available from the manufacturers of the fluorometers and the U.S. Geological Survey (1986, 1989). For a mill with multiple discharges of process water, it is recommended that each discharge be dye-traced separately to define its individual effluent plume configuration.

The recommended method is as follows:

i. Carefully record the location of the tracer concentration measurements in the plume. The precision of positioning systems would be determined on a site-specific basis.

ii. Measure the tracer concentrations along lines perpendicular to the direction of travel of the plume by traversing the plume in a boat equipped with a recording flow-through fluorometer and a pump sampling system, or preferably an *in situ* submersible fluorometer. A towed *in situ* fluorometer used in conjunction with an accurate real-time positioning system is recommended. Such a system can be repeatedly towed through the water column to rapidly build a three-dimensional picture of the effluent concentration.

In situ fluorometers overcome many of the problems of a cheaper pumped sampling system, such as, mixing within the sampling hose, time delay caused by travel through the sampling hose, and the resulting slow response to concentration gradients. When the dye concentration profiles are measured *in situ* there is no time delay between the sample measurements and the location of the boat, compared to using a pumped sampling system. Grab samples are recommended only if the plume is not accessible by boat. Grab samples are collected for subsequent analysis in the laboratory. At least 12 samples at each transect are recommended in the plume to adequately define the plume configuration. This procedure is slower, has poor spatial resolution, and does not provide a continuous profile of the tracer concentrations. Consequently, it is difficult to carry out a mass balance of the tracer.

iii. The depth of sampling should match the depth distribution of the tracer. The depth of maximum tracer concentration (trapping depth) is used for the concentration profiling of the tracer. By using an effluent tracer, the effects of stratification in the receiving water on the effluent are directly measured. Where the effluent is not well mixed with depth, the vertical profile should be defined. Towing should be done at the maximum concentration depth with periodic checks to ensure towing is at maximum concentration.

iv. Sample the effluent plume at various distances downstream, approximately doubling the downstream distance intervals between transects to identify unusual characteristics in the plume. Unusual characteristics could include high concentrations observed at places beyond where the effluent concentration has dropped below the specified concentration (e.g. accumulation in an inlet, bay, etc.), or where an undertow moves effluent downstream or off-shore and resurfaces elsewhere, or similar situations.

v. Measure and record the bathymetry (depth sonar techniques are generally adequate) at tracer profiling locations. Determine positions of the transects using a locating system such as satellite fixing or optical fixing on shore landmarks. Where a detailed hydrographic chart is already available, a survey of bathymetry may not be necessary. It is recommended that the bathymetry data of the receiving water be presented on a map of the exposure area.

vi. If the effluent plume is influenced by other effluent plumes in the vicinity, design the water quality sampling program to consider these other plumes. For other discharges upstream, collect a water quality sample upstream from the target discharge and treat it as a background concentration for the effluent plume water quality.

2.2.3 Using Predictive Models

The use of predictive plume models allows for a more complete characterization of the plume based on samples collected during the plume delineation study. Modelling provides a means of extrapolating from the plume measurements to define the extent of far-field conditions and to simulate the continuous injection of effluent typical of site operating conditions. Plume models also assist in the interpretation of collected data when the plume configuration is strongly influenced by the direction of ambient currents (e.g. lakes and oceans) and when the receiving body is characterized by a long flushing time (e.g. stagnant waters).

When ambient currents modify plume configuration, the plume's characteristics can be established through statistical analysis of the data collected by the current meters. For conditions of a long flushing period, the measured dye concentrations are used to calibrate a numerical transport-diffusion model. The model can then be used to simulate the effluent delineation and characteristics arising from a continuous discharge. The model can be run for a variety of conditions (e.g. seasonal variations of water movements and wind patterns), thereby overcoming the limitations of the particular conditions recorded in a single field study. The output from the model can be statistically analyzed to estimate the desired regions of maximum extent, average conditions, and minimum dilution. Discussion of appropriate plume models capable of such calculations can be found in U.S. Army Corps of Engineers (1986), and in Liu and Leendertse (1987).

2.2.4 Environment-Specific Considerations

For the purposes of this document, the type of receiving water is based on the most predominant characteristics as follows:

river - current in one direction only, flow seasonally variable
lake - currents in many directions
estuary - stratified fresh water and salt water with tides
fjord - relatively small fresh water discharges with tides
ocean - sea water with tides and currents in many directions

Recommended methods for different receiving water types follow.

2.2.4.1 Rivers

To bracket the range of expected conditions, plume delineation studies should be carried out both during the annual low and maximum river flows. If one of these conditions occurs under an ice cover, the lowest or maximum annual river flow in the ice-free period should be used given operational and safety considerations. The continuous flow constant rate tracer injection technique should be used (Section 2.2.2.4). The number of measurements will be site-specific. It is recommended that concentrations be measured at various distances downstream and that the spatial extent of the plume be determined. Slug tests are also a suitable technique for monitoring mixing in rivers.

River flow for long-term average conditions can be interpolated from measured plume configurations under the annual low and high river flow conditions (U.S. EPA 1987). Calculations and modelling can be used for this interpolation.

2.2.4.2 Lakes

The recommended method for the delineation of the effluent plume for a lake discharge is the constant tracer injection procedure. The slug test is also a suitable technique for monitoring mixing in lakes. General considerations for designing the study are as follows:

i. The direction of movement and dilution characteristics of the plume are variable depending on the prevailing currents.

ii. The effluent may not be well mixed with depth.

iii. Most of the time the effluent plume for a mill discharge is near the water surface.

iv. The tracer has the same density as the effluent.

The tracer studies should be designed to fulfil the general requirements and to consider seasonal variations. The duration and timing of each study will be influenced by prevailing currents.

It is recommended that drogues be released and tracked concurrent with the dye tracing, to measure currents. The drogue water sails should be set 1.5 m below the water surface or set at the trapping depth of the effluent, which can be determined on site from the initial tracer monitoring. If clusters of drogues are released, dispersion for a batch release can be obtained from the paths of the individual drogues by determining the variance of the individual drogues about the centroid.

It is recommended that at least two recording current meters of appropriate sensitivity and a wind anemometer be operated for a period of a month for each season of measurement to monitor for seasonal variation in the vicinity of the outfall. Data from the recording current meters are analyzed to quantify the statistical characteristics of the current climate in the vicinity of the outfall. Site-specific evaluations of local airport or other weather data should be done to establish the usefulness of such data. The statistical relationship between the dilution and travel time is then combined with the statistical characteristics of the currents to develop spatial dilution envelopes (a graph showing the frequency of dilution factor as a function of two variables: dilution and travel time) about the outfall.

The dilution envelopes should be developed and statistical confidence determined and recorded. These envelopes should show the probability of the effluent plume being present at any receiving water location and the mean and standard deviation of the effluent concentration at any location in the receiving water. The envelopes do not show the configuration of the effluent plume under specific current speed and direction conditions.

To determine the effluent plume configuration at any particular time, and hence for a specific condition of current speed and direction, it is necessary to carry out a tracer study.

Unlike the river receiving water where the plume configuration can be predicted for the design condition using models, the design conditions for lakes are not easily predicted (see Section 2.2.3). Some provincial regulatory agencies specify conditions for which the effluent plume should be defined. For example, in Ontario, the Ministry of the Environment and Energy requires that the plume be defined for the high velocity condition normally greater than 20 $cm \cdot sec^{-1}$ in both directions parallel to shore and the stagnation condition.

2.2.4.3 Estuaries

In most instances, industrial sites are located at the freshwater end of an estuary, primarily because the site requires process water from a freshwater source. There are exceptions where sites obtain their freshwater from other sources or small streams. In these instances, the ocean processes predominate and require special studies. These are discussed in the following as fjords.

In estuaries, in most cases, tidal effects and the depth distribution of the effluent, which is fresh water in salt water, are important characteristics. The currents in estuary receiving waters are to a large extent determined by the tidal variations.

Over the range of tidal conditions, tidal velocities are largest during the rising (flooding) and falling (ebbing) tides for the spring tidal condition. The effluent plume should be measured during maximum flood and ebb currents and at slack tide to delineate the minimum and maximum extent of the effluent plume. The rising and falling tide studies should be carried out during the period when river flow into the estuary is at an annual low.

The suggested method for estuary tracer studies is the continuous constant rate injection of a tracer for a period of about 4 hours starting at the beginning of the rising or falling tide. Mix the tracer with the effluent before discharge and measure the resulting tracer plume at right angles to the direction of movement of the plume and at various depths. To completely sample the plume, it may be necessary to carry out tracer injections over several tidal cycles, allowing sufficient time between injections to flush the previous tracer out of the exposure area. Salinity-temperature-depth (STD) data should also be collected in each of the tracer profile cross-sections. Records of tidal stage and currents in the areas affected by the plume should be kept during the tracer studies.

The measured plumes for the rising and falling tides represent the maximum upstream and downstream extent of the plume. These plumes are plotted as percentage effluent. The slack tide conditions are between the rising and falling tides and can be obtained by drawing envelopes about the data collected from this period.

2.2.4.4 Fjords

In a fjord, plume transport and dilution are caused primarily by residual currents. Tidal activity, while present, does not contribute to net effluent movement, but does affect effluent dispersion. The residual current pattern can be determined in the vicinity of the discharge by injecting tracers into the effluent at a continuous rate over three tidal cycles, and measuring the resulting dye plume during slack tide after the second and third tidal cycles.

Concurrent with the tracer studies, at least two recording current meters operated at maximum concentration depth in the fjord are recommended during the study and for an additional 60-day period after completion of the study. Using the statistical data on the residual currents and measured tracer plume data, dilution envelopes can be developed.

Some fjords are very complex, and using locally measured plume characteristics and residual currents may not be sufficient to delineate the effluent plume, particularly for the far-field or for low effluent concentrations. In these cases it may be useful to combine a residual current transport model with a contaminant model such as the WASP (U.S. EPA 1988).

Normally, carrying out the studies during the spring rising and falling tides when the river flow is at its annual low flow is adequate for design conditions. Prediction models may be useful to determine the plume. Models for predicting the effluent plume for spring conditions are discussed in U.S. Army Corps of Engineers (1986) report on the CE-QUAL-W2 model.

2.2.4.5 Oceans

Ocean receiving waters are distinguished from estuaries in that there is no fresh water in the vicinity of the discharge. The effluent plume is diluted and transported by currents generated by tides, residual currents, and winds. The recommended method is to continuously inject a tracer into the effluent for three tidal cycles (under spring tide conditions), and to measure the plume during the slack current after the second and third tidal cycles. Tracer concentration profiles collected in the plume in two directions set at right angles are recommended. The depth of the profiles is determined by measuring the tracer concentration depth profiles in the plume, then selecting the depth of the maximum concentration as the depth for the tracer concentration profiles.

At least two recording current meters at the maximum concentration depth and a wind anemometer are recommended during the tracer studies and for at least a 60-day period after the studies in the vicinity of the discharge.

From the tracer profiles, the effluent plume envelopes are developed and are related to the measured currents. The probability of the occurrence of these envelopes is determined from statistical analysis of the current meter records.

2.3 Habitat Inventory and Classification

The classification of the fisheries habitat that may potentially be affected by mill discharges or confounding discharges is critical to the EEM design. The recommended method to present habitat information is to map all relevant areas as follows:

1) areas delineated as zone of effluent mixing determined in Section 2.2.1;
2) areas in or proximate to this zone that are potentially affected by deposition of sediments contributed by the mill or contaminated by mill effluent; and,
3) areas within the reference area(s) with similar dominant habitat types to those identified in (1) and (2) above.

The features to be identified and mapped as part of this inventory are described in Table 2.2. This will assist in the identification of potentially confounding factors to be considered in the design of the study and interpretation of results such as dams, tributaries, other discharges, etc. The spatial extent of the study area to be mapped will be determined on a site-specific basis. However, since sampling areas for invertebrate surveys must be located in the dominant habitat class (as defined in Cowardin *et al.* 1979), all habitat areas to be mapped must be at least to the level of class.

The habitat classification studies must follow the hierarchical classification system of Cowardin *et al.* (1979). Additional guidance for habitat description procedures is also available in Department of Fisheries and Oceans (1990).

Information from habitat inventory and classification is used to locate sampling stations within similar habitats in the exposure and reference areas, and to obtain information on other discharges and modifying factors that could confound the interpretation of data obtained from those locations.

Table 2.2 Reporting requirements for resource and habitat inventory of the study area

Parameter	Description/Information to be Reported
Major tributaries, river mouth	• locations to be shown on maps
Fish spawning grounds, nursery areas	• locations of all known areas to be shown on map
Fishing grounds, aquaculture areas, significant shellfish resources	• locations of all known areas to be shown on map
Water intakes, effluent discharges, stormwater discharges, sewer overflows, disposal sites, log booming and storage areas	• locations of all known areas to be shown on map
Docks, wharves, ferry terminals, marinas, boat launches, and public beaches	• locations of all known areas to be shown on map
Dams and other barriers to fish	• locations of all known areas to be shown on map and fish movement restrictions indicated
Zones of plant growth	• identify any areas in near-field plume where aquatic plant growth appears to be reduced or enhanced relative to reference areas.
Bathymetry mapping	• units to be reported in meters (relative to mean sea level in tidal habitats) • recommended intervals: ◆ 0-5 m water depth - 1 m intervals ◆ 5-10 m water depth - 2.5 m intervals ◆ >10 m water depth - 5 m intervals
Gradient (rivers)	• gradient profile for riverine study areas, produced from topographic maps.
Discharge (rivers and estuaries)	• summary water discharges statistics (m^3/s), as available for each month of record (gauged watersheds) for most representative station(s); • substantiated estimate of mean monthly discharge when data unavailable • discharge during previous EEM cycle
Water chemistry	• table of available historical water quality data
Water currents (lakes, estuaries and oceans)	• compile and map available data
Tides (marine)	• tabulate available tide height data from government tide tables for nearest station

Note: recommended mapping scale is 1: $\leq$ 5,000 and maps should include all areas that may be included in the exposure and reference areas

2.3.1 Environment-Specific Considerations

Rivers

It is recommended that river habitat descriptions include information on elevation gradient; the location of dams, falls and other barriers to fish migration; mean annual discharge and ranges; and general substrate characteristics of each river (preferably in the form of a gradient profile chart). Upstream and downstream inputs (e.g. storm water, sewer overflow, effluent from other industrial sites) should be mapped and described.

Lakes

Important habitat features of lakes include bathymetry, the locations of major inlets and outlets, and general oxygen-temperature conditions (e.g. thermal stratification, occurrences of oxygen depletion in deep water).

Open Coastlines

Suggested additional mapping parameters for open coastlines (marine, Great Lakes) include depth contours, near-shore substrate characteristics, shoreline configuration, and the locations of inflowing rivers and other discharges and activities.

Estuaries

Estuaries are best described in terms of their general salinity gradients, flows, bathymetries, and general substrate features. A description of tidal cycles is recommended for all marine and estuary situations. Most of the above features can be described from navigational maps, topographic maps, government publications on tides and river discharge records, and through interviews with local government officials and knowledgeable individuals.

2.3.2 Habitat Classification

The recommended framework for classifying and mapping aquatic features is the classification system developed by the U.S. Fish and Wildlife Service (FWS), "Classification of Wetlands and Deepwater Habitats of the United States" (Cowardin *et al.* 1979, Busch and Sly 1992). The system allows for classification of a wide range of continental, aquatic, and semi-aquatic habitats.

The FWS classification system is hierarchical. Habitats are divided among five major systems at the broadest level: marine, estuarine, riverine, lacustrine, and littoral (Figure 2.1). Systems are further subdivided into subsystems that reflect hydrologic conditions, e.g. littoral vs. limnetic in the lacustrine system. However, the littoral system is not divided into subsystems. Below the subsystem level is the class level, which describes the appearance of the habitat in terms of vegetation (e.g. emergent wetland, aquatic bed, forested wetlands) or in terms of substrate where vegetation is inconspicuous or absent (e.g. unconsolidated shore, rocky shore, streambed). Each class can be further subdivided into subclasses and subclasses divided into dominance types.

For the purposes of EEM, it is recommended that the classification system be used down to the class level, as shown in Figure 2.1. In addition, the FWS Subclass and Dominance Types for substrate and modifiers should be used for salinity and water regimes (Tables 2.3 to 2.5). Because the FWS system provides little or no resolution in the vertical (depth) dimension, collection of information on temperature and oxygen profiles during limiting times of the year (generally summer or possibly under ice in winter), and both vertical and horizontal salinity gradients under various tidal conditions is recommended if the receiving system is subject to vertical stratification (e.g. estuarine).

Where the plume is very large and complete mapping of subclasses and dominance types for substrate is impractical, detailed classification in the vicinity of effluent discharge (in the initial discharge zone and near-field plume area) for effluent dilutions higher than the specified concentration is suggested. Additional guidance on habitat description procedures for coastal and estuarine situations is provided in Cowardin *et al.* (1979). In addition, it is recommended that bottom substrates (where appropriate) be described in relevant industry specific parameters. For example, for pulp and paper use the degree (thickness) of deposition of fibre and woody debris.

Further guidance on aquatic habitat assessment can be found in the Department of Fisheries and Oceans and the British Columbia Ministry of the Environment and Parks (1987), Orth (1989), Ontario Ministry of Natural Resources (1989), Plafkin *et al.* (1989), and the Department of Fisheries and Oceans (1990).

2.3.3 Delineation of Depositional Zones

Depositional zones occur when water velocity decreases resulting in particles settling out; the finest particles settle out in the slowest current speeds. Historical contaminant or benthic community data, in conjunction with the effluent plume delineation results, may be helpful in identifying exposure areas. To contrast resident benthic invertebrate communities, similar (but uncontaminated) sediment depositional zones should be located in the reference area.

Figure 2.1 Classification hierarchy of wetlands and deep water habitats showing systems, subsystems, and classes (after Cowardin *et al.* 1979)

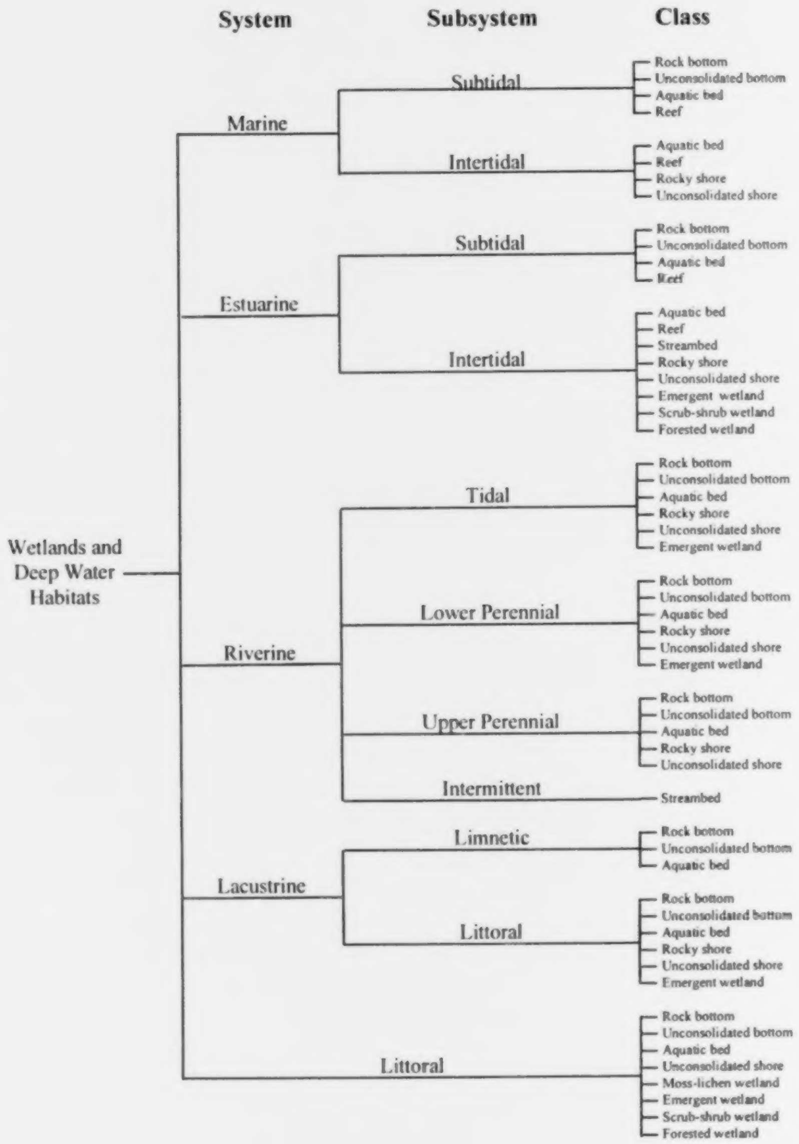

System
Subsystem
Class
Wetlands and Deep Water Habitats
Marine
Subtidal
Rock bottom
Unconsolidated bottom
Aquatic bed
Reef
Intertidal
Aquatic bed
Reef
Rocky shore
Unconsolidated shore
Estuarine
Subtidal
Rock bottom
Unconsolidated bottom
Aquatic bed
Reef
Intertidal
Aquatic bed
Reef
Streambed
Rocky shore
Unconsolidated shore
Emergent wetland
Scrub-shrub wetland
Forested wetland
Riverine
Tidal
Rock bottom
Unconsolidated bottom
Aquatic bed
Rocky shore
Unconsolidated shore
Emergent wetland
Lower Perennial
Rock bottom
Unconsolidated bottom
Aquatic bed
Rocky shore
Unconsolidated shore
Emergent wetland
Upper Perennial
Rock bottom
Unconsolidated bottom
Aquatic bed
Rocky shore
Unconsolidated shore
Intermittent
Streambed
Lacustrine
Limnetic
Rock bottom
Unconsolidated bottom
Aquatic bed
Littoral
Rock bottom
Unconsolidated bottom
Aquatic bed
Rocky shore
Unconsolidated shore
Emergent wetland
Littoral
Rock bottom
Unconsolidated bottom
Aquatic bed
Unconsolidated shore
Moss-lichen wetland
Emergent wetland
Scrub-shrub wetland
Forested wetland

Table 2.3 Substrate subclasses and dominance types (after Cowardin *et al.* 1979)

Substrate Class	Subclass/Dominance Type	Substrate Class	Subclass/Dominance Type
Rock bottom	bedrock rubble	Unconsolidated bottom	cobble-gravel sand mud organic
Aquatic bed	algal aquatic moss rooted vascular plants floating vascular plants	Reef	mollusc
Streambed	bedrock rubble cobble-gravel sand mud organic	Rocky shore	bedrock rubble

Table 2.4 Salinity modifiers (after Cowardin *et al.* 1979)

Modifiers	Salinity (parts per thousand)	Approximate Specific Conductance (µmhos at 25°C)
Hyperhaline	>40	>60,000
Euhaline	30-40	45,000-60,000
Mixohaline (Brackish)	0.5-30	800-45,000
Polyhaline	18.0-30	30,000-45,000
Mesohaline	5.0-18	8,000-30,000
Oligohaline	0.5-5	800-8,000
Fresh	<0.5	<800

Table 2.5 Water regime modifiers (after Cowardin *et al.* 1979)

Subclass	Modifier	Subclass	Modifier
Tidal	Subtidal	Nontidal	Permanently flooded
	Irregularly exposed		Intermittently exposed
	Regularly flooded		Semi-permanently flooded
	Irregularly flooded		Seasonally flooded
			Temporarily flooded
			Intermittently flooded

It is recommended that depositional zones adjacent to the outfall (near-field) and further down current (far-field) be mapped and characterized. The spatial extent and depth of these deposits can be determined in a site-specific sampling survey using methods such as sediment grabs, cores, dredges, diver-deployed sediment traps, or other appropriate methods. Depth of deposits can be measured directly, where possible, or estimated using physical probing, sonar, or geophysical methods. Sampler contents can be described in terms of sediment layer texture, colour, fauna, and debris. A colour photographic record of the sample is also useful.

2.4 Resource Inventory

The identification of aquatic resources and resource uses which may potentially be affected by the effluent discharge is critical to the design of the EEM study and the selection of appropriate sentinel species for the fish survey. As part of the pre-design requirements for EEM, description of fisheries resources within the study area is required. This information should include the identification of fish and shellfish that are presently exploited in commercial and non-commercial fishing, that may potentially be exploited in the future, and any species recognized by federal or provincial authorities as rare, threatened or endangered, as well as any species which may be present in sufficient numbers to be considered as a monitoring species. Aquaculture installations should also be noted. This information can usually be obtained from district fisheries biologists in federal or provincial regulatory or museum agencies, from local conservation officials and from interviews with fishermen and members of the local community. The mills must be able to demonstrate to the RAO that a reasonable level of effort has been expended to gather this information.

While mills are not required to undertake a preliminary field survey in order to collect fisheries resource information, they must have sufficient site-specific data to allow selection of sentinel species of fish (or shellfish) that satisfy the criteria for the fish survey (see Section 5). This would normally be done by reviewing the local fish resource species information for candidate species as described above, or consulting reference literature on fish distributions in Canada. In certain instances, this may require undertaking a non-destructive preliminary fish survey to document species presence and relative abundance. Mills with knowledge of fisheries resources in the study area are generally more successful in capturing sentinel species.

The reporting requirements for resource and habitat inventory are listed in Table 2.2. All habitat classification is to be reported in map form, typically at a scale of $1{:}\leq 5{,}000$ (actual scale to be reported). A suggested format for reporting resource and sentinel species information is given in Table 2.6.

Table 2.6 Example of resource inventory information

Resource species and relative abundance (if known)				
Commercial	Sport	Subsistence	Rare/Endangered	Others
Candidate sentinel fish species and habitat references				
Species	Spawning	Nursery	Forage	Migration
Fish community assessment reports (indicate if none exist)				
Author	Title	Citation	Conclusions	
Fish contaminant reports (indicate if none exist)				
Author	Title	Citation	Conclusions	

2.5 Detailed Site Map

Preparation of a site map that illustrates the estimated extent of known locations of sensitive resource/habitat values, and locations of previous sample stations is recommended. A site map is useful to illustrate the consolidation of available data and substantiate the selection of future sampling stations and monitoring variables.

2.6 Historical Receiving Environment Data

The review of historical (10 years) and more recent monitoring data which has been collected for other purposes is an important component of the pre-design information requirements. This information may assist in identifying known impacts, if any, and will aid in determining selection of exposed and reference areas. This information or data may be provided as part of the pre-design stage if the data are of a sufficient quality to meet the requirements of the EEM program. This review should also identify other site-specific monitoring programs, e.g., intensive dissolved oxygen monitoring programs, which should be undertaken or are being conducted on a site-specific basis. This review should also identify past problems or concerns, such as tainting, shellfish or fisheries closures, bacterial contamination and sediment fibre mats.

2.7 Effluent Quality and Characterization

The *Pulp and Paper Effluent Regulations* under the *Fisheries Act* and the *Regulations Respecting the Release of Polychlorinated Dibenzo-para-dioxins and Polychlorinated Dibenzofurans in Effluent from Pulp and Paper Mills* under the *Canadian Environmental Protection Act* require that mills report on effluent quality. Effluent should be characterized for following parameters:

- pH
- flow
- conductivity
- biochemical oxygen demand (BOD)
- total suspended solids (TSS)
- 96-h rainbow trout lethality
- 48-h *Daphnia magna* lethality
- chlorinated dioxin and furan congeners[5]

Table 2.7 provides an example of how to summarize effluent quality and toxicity data on an annual basis. Data can also be presented graphically. Substantial changes in processes or effluent treatment and data which reflect current mill operating conditions can be noted.

---

[5] Requirements for *Regulations Respecting the Release of Polychlorinated Dibenzo-para-Dioxins and Polychlorinated Dibenzofurans from Pulp and Paper Mills* (*Canadian Environmental Protection Act*)

Table 2.7 Example of annual summary of mill effluent characteristics[1]

Chemistry data[2]

Year	Mean daily production	Mean minimum flow (m^3)	Mean daily flow (m^3)	Maximum flow (m^3)	Number of samples analyzed	Concentration range	Mean concentration of samples above MDL[3]	Standard deviation	Comments (changes in process effluent treatment)

Toxicity data[4]

Year	Mean daily production	Mean daily flow (m^3)	Number of samples analyzed	Endpoint measured	Endpoint range	Mean endpoint samples	Standard deviation	Comments (changes in process effluent treatment)

[1] Complete separate forms for each effluent discharge.

[2] Fill in separately for each chemical analyzed.

[3] Censored data (those associated with "<" or ">" as in <0.05 mg · L^{-1} or >100% effluent) should be incorporated into mean as limit of quantification (e.g. 0.05 or 100, corresponding to above examples), however, mean presented in table should be flagged as to proportion of data that were censored.

[4] Fill in separately for each toxicity test type conducted.

2.8 Mill History and Operations

Relevant historical data regarding mill history, especially processing, effluent treatment and spills should be reported, as they may affect interpretation of study design or results. This information is critical for identifying acceptability of historical data, and for selecting study areas. A review of historical and current mill operational data may also enable a mill operator to identify environmental concerns that can be attributed to current or past practices. An example of a checklist of specific items that could be documented for a pulp and paper mill is presented in Table 2.8. To provide a better understanding of current operations, a simple schematic diagram of major sewer connections and corresponding flows of a mill could be drafted.

Table 2.8 Example of pulp and paper mill operations checklist (check when complete)[1]

Documentation	Historical operations	Current operations	Significant changes since previous EEM[2]
Process and Non-Process Water Sources			
Effluent Discharge Recipient			
Pulping Process:			
• general description			
• describe chemicals used which ultimately enter sewer system			
• volume and rate of use of all chemicals ultimately discharged to collection sewers			
Description of Furnish:			
• type (round wood, chips, other)			
• species			
• volumes and age of each type and species (hardwood, softwood)			
Bleaching Process:			
• general description			
• bleaching agent(s)			
• volume and rate of use of all chemicals ultimately discharged			
Treatment Facilities:			
• general description including retention time			
• volume and rate of any chemicals added			
• spill or equalization pond and residence time			
Losses (fibre, brownstock washer, soda)			
• quantities			
• types			
• dates			
Schematic Diagram Showing:			
• location and flow of all major process and cooling water sewers			
• location of treatment facilities			
• location of flow measuring devices			
• landfill and leachate collection information			
• surface runoff discharge points			
Flow Measurement Devices:			
• location and type			
• dates calibrated			
Effluent Monitoring Data			
Spills to Receiver[3] (type, volumes, dates)			

[1] Supporting documentation should be provided in EEM report.

[2] Document all significant changes and dates of change.

[3] Document all spills which demonstrated, or had potential to cause, environmental impact.

2.9 References

Busch, W.-D. N. and P. G. Sly. (eds.) 1992. The development of an aquatic habitat classification system for lakes. CRC Press. Boca Raton. 225 pp

Cowardin, L.M., V. Carter, F.C. Golet and E.T. LaRoe. 1979. Classification of Wetlands and Deepwater Habitats of the United States. U.S. Fish and Wildlife Service. FWS/OBS-79/31.

Department of Fisheries and Oceans. 1990. Coastal/Estuarine Habitat Description and Assessment Manual. Part II. Habitat description procedures. Prepared by G.L. Williams and Associates Ltd., Coquitlam, B.C.

Department of Fisheries and Oceans and British Columbia Ministry of the Environment and Parks. 1987. Fish Habitat Inventory and Information Program: Stream Survey Field Guide.

Douglas G.R., C.E. Grant, R.D.L. Grant, M.F. Salamone and E.R. Nessman. 1983. Comparative mammalian in vitro and in vivo studies on mutagenic activity of Rhodamine WT. Mutat. Res. 118: 117-125.

Liu S.K. and J.J. Leendertse. 1987. Modelling the Alaskan Continental Shelf Waters. The Rand Corporation, 1700 Main Street, P.O.Box 2138, Santa Monica, California. R-3567-NOAA/RC.

Ontario Ministry of Natural Resources. 1989. Manual of Instructions. Aquatic Habitat Inventory Surveys. MNR, Toronto, Ontario.

Orth, D.J. 1989. Aquatic habitat measurements. In: Fisheries Techniques, L.A. Neilson and D.L. Johnson (eds.), pp. 61-84. Am. Fish. Soc., Bethesda, Maryland.

Parker, G.G. 1973. Tests of Rhodomine WT dye for toxicity to oysters and fish. J. Res. U.S. Geological Survey 1(4): 499.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross and R.M. Hughs. 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and fish. EPA/444/4-89-001.

U.S. Army Corps of Engineers. 1986. CE-QUAL-W2: a Numerical Two-Dimensional, Laterally Averaged Model of Hydrodynamics and Water Quality; User's Manual. Instruction Report E-86-5, U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi. U.S. EPA. 1987. The Enhanced Stream Water Quality Models QUAL2E and

QUAL2E - UNCAS: Documentation and User Manual. Environmental Research Laboratory, United States, Athens, Georgia. EPA/600/3-87/007.

U.S. EPA. 1987. The Enhanced Stream Water Quality Models QUAL2E and QUAL2E - UNCAS: Documentation and User Manual. Environmental Research Laboratory, United States, Athens, Georgia. EPA/600/3-87/007.

U.S. EPA. 1988. WASP-4 a Hydrodynamic and Water Quality Model-theory, User's Manual and Programmer's Guide. Environmental Protection Agency, Athens, Georgia. EPA/600/3-87/039.

U.S. Geological Survey. 1986. Fluorometric Procedures for Dye Tracing. U.S. Geological Survey, Federal Center, Box 25425, Denver. Book 3. Applications of Hydraulics. Chapter A12.

U.S. Geological Survey. 1989. Determination of Stream Reaeration Coefficients by Use of Tracers. U.S. Geological Survey, Federal Center, Box 25425, Denver. Book 3. Applications of Hydraulics. Chapter A18.

3 Sampling Design

This general chapter on design is applicable to all EEM field monitoring activities. Some examples for benthos or fish are given for clarity. Details on specific sampling designs related to fish and benthos components are found in sections 5 and 6, respectively.

3.1 Description of Study Areas

An area is qualitatively defined for sampling purposes and relates to the appropriate geographical scale encompassing one or more fundamental sampling locations called stations. A station is a fixed sampling location which can be recognized, re-sampled and defined quantitatively (e.g. latitude /longitude). Within EEM, the overall study area is sub-divided into reference and exposure areas (near-field, far-field, etc.). Also, area is defined to be relatively homogeneous with respect to major habitat class and level of exposure to mill discharge.

A reference area should have zero mill effluent exposure and have natural habitat features similar to those of exposure areas. In most cases, the boundary of the exposure area is defined by the zone of effluent mixing. Within an exposure area, there will be near-field and far-field areas. Near-field areas are outside the initial discharge zone and have higher exposure to effluent than far-field areas. At least one of the near-field areas should be as near as possible to the effluent discharge point but located outside the initial discharge zone. Far-field areas are outside the zone of any acute lethal effects results. Far-field areas are recommended to be positioned close to the boundary of the zone of effluent mixing. Multiple sampling stations in each defined area should be used to determine spatial variation. In a gradient design, there is no reference area per se, but the response variables are evaluated along the exposure gradient.

In practical use, there will probably always be one or more far-field areas for media other than fish (e.g. water, sediment, and benthos). Recommended positioning of far-field areas should include variation of exposure to the effluent. All far-field areas should be located, if possible, to minimize or to avoid exposure to non-mill discharges.

Most effluents are discharged so that the velocity of the effluent exceeds the velocity of the receiving water. The area where this occurs and the effluent is buoyant is referred to as the *initial discharge zone*. The initial discharge zone is often characterized by visual turbulence and typically does not extend more than 5 to 50 metres from the outfall.

Reference areas should have natural habitat features similar to those in exposure areas.

Separate reference areas should be selected for each habitat type identified in the exposure area. Comparisons are only made within major natural habitat categories. Considerations for selection of reference areas in marine, estuarine, and freshwater environments are presented in Sections 3.2 and 3.3.

For mills with insufficient historical information to locate reference and exposure areas, exploratory sampling may be useful. Exploratory sampling can also be used to identify habitat characteristics for effective selection of sampling stations.

An experienced field crew can approximate the effluent field based on field measurements of specific conductance or preliminary dye study results, and can often identify likely depositional areas based on observed receiving water flow and circulation patterns. Thus, it is usually possible to choose some appropriate water and sediment sampling stations in the field and to complete exploratory sampling of the receiving environment concurrent with plume and depositional zone studies and critical resource/habitat inventories in a single campaign.

3.2 Selection of Reference Areas in Marine and Estuarine Environments

In order to determine the most appropriate sampling design for a mill, it is necessary to determine if suitable reference areas are available upstream or within the same water body or ecoregion as the mill. This is possible if physically and biologically comparable stations can be found inside and outside the zone of effluent exposure and the assumption can be made that multi-species organism abundance at these treatment and control stations respond similarly to environmental conditions. If suitable reference sites can be found, then control/ impact type sampling designs may be used. These criteria emphasize that the reference area must be suitable physically as well as biotically, and outside the influence of the mill or other confounding factors. Usually, coastal mills do not have strictly "upstream" areas for reference sampling because of variable directionality of current due to tidal effects. Estuarine mills may have upstream areas which are too different physically and biotically to be suitable for reference sampling. The reference area is, therefore, usually at least periodically downstream from the effluent discharge. Therefore it is important to understand the current flow patterns in the area to determine whether a potential reference area is "outside" mill effluent exposure. In section 6.1.5, some criteria are given for determining the suitability of reference sites, and to help decide if reference areas can be identified.

3.3 Selection of Reference Areas in Freshwater Environments

Environmental and biological data obtained from reference areas when compared to exposure areas detect impairment of aquatic life (Yoder 1991), diagnose stressors (Hughes *et al.* 1994), provide data on temporal and spatial trends (Yoder 1989) and provide data for water resource summaries for government agencies (OEPA 1990). Reference areas are not to represent pristine (or pre-European settlement) conditions, but areas in which impacts are lowest or disturbance is minimal (Simon 1991, Omernik 1995).

The identification of "least impacted" areas will differ across the country. For example, minimally impacted reference areas in the prairies may be rivers in drainage basins that lack point source effluents, logging activities, road building or cattle grazing. The areas, however, will have a high degree of agricultural activity. If sufficient riparian vegetation has been allowed to grow to provide a buffer zone between the land and receiving waters, these areas would be suitable reference areas. Reference areas in extremely disturbed areas may be impossible to locate. Here, schemes must be devised so that reference areas with minimal degradation are located in comparable drainage basins within the same ecoregion (Hughes *et al.* 1994).

3.3.1 The Control-Impact (C-I) Approach in Reference Areas Selection

The simplest control-impact (or reference-exposure) design (Green 1979) used for environmental effects monitoring in rivers consists of one upstream reference area and a series of downstream exposure areas representing near-field, far-field and far far-field. A similar situation exists for mills located on lakeshores. One reference area is typically selected in a nearby bay of the same lake and exposure areas are selected at increasing distances away from the mill effluent.

Single reference areas are criticized for their pseudo-replication (Hurlbert 1984). An upstream reference site may differ from an exposure site with respect to many variables (substrate, current velocity, environmental gradients, river valley, land use) in addition to the point source. Thus, the problem with a single reference is that site selection is often confounded (Resh 1995). Moreover, it is not uncommon for populations in "two locations to diverge or converge through time" in the absence of an environmental impact (Underwood 1991). Clearly, a single reference or control site may not be adequate to evaluate the effect of a potential impact because it does not provide sufficient information on natural variability.

3.3.2 Classification Scheme for Reference Area Selection

Because reference areas will vary among different landscapes, approaches have been developed to classify the land that rivers run through or in which lakes reside to predict aquatic biotic assemblages (Corkum 1989, 1992; Hughes 1995, Maxwell *et al.* 1995, Omernik 1995). A classification system is a way of simplifying sampling procedures and management strategies by organizing a variable landscape (Conquest *et al.* 1994). The assumption is that the classification scheme is hierarchical. The advantage of a hierarchical classification scheme is that it "offers a way to discriminate among features of the landscape at several scales of resolution" (Conquest *et al.* 1994). The classification scheme is based (with modifications) on one developed by the U.S. Department of Agriculture Forest Service (Maxwell *et al.* 1995) (Figure 3.1).

The hierarchical classification scheme is presented as a guide in the *a priori* selection of reference areas. If mill effluents are located on rivers where upstream inputs are absent, reference areas may be selected within the drainage basin in which the mill is located. If non-point or other point source inputs occur upstream of an effluent, reference areas should be selected from the nearest comparable drainage with minimal development. Drainage basins selected must occur within the same ecoregion in which the mill site is located.

In lakes with a single mill effluent and without non-point sources of pollution, the sphere of influence originating from the effluent should be determined. This is particularly important for lakes in which effluent flow is not unidirectional. If tracer studies indicate that effluent effects are restricted, reference areas may be selected within the lake in which the mill occurs. These reference areas must occur in separate, but comparable bays or basins of the lake. If tracer studies reveal that the identified effluent is dispersed throughout the lake or if other sources of pollution exist, reference areas should be selected in the nearest comparable lake or lakes within the same or adjacent drainage basin and within the same ecoregion in which the mill site is selected.

3.3.2.1 Ecoregions

The first step in reference site selection is to use terrestrial attributes (ecoregions) with similar features. Ecoregions are defined as "part of an ecoprovince characterized by distinctive ecological responses to climate as expressed by vegetation, soils, water, and fauna" (Wiken 1986, Wickware and Rubec 1989). Ecoregion maps for Canada are available on the Internet (Figure 3.2). The WEB site address is:

http://www.cciw.ca/eman-temp/ecozones/intro.html.

Figure 3.1 Classification scheme for reference site selection

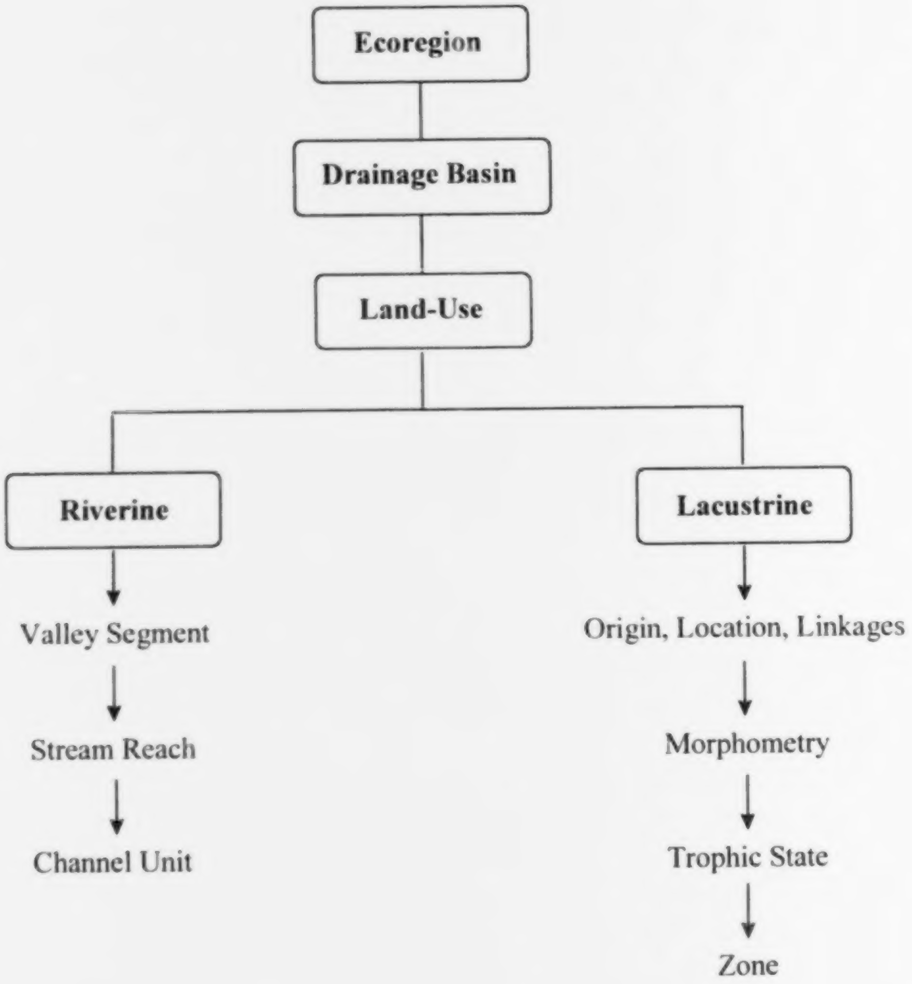
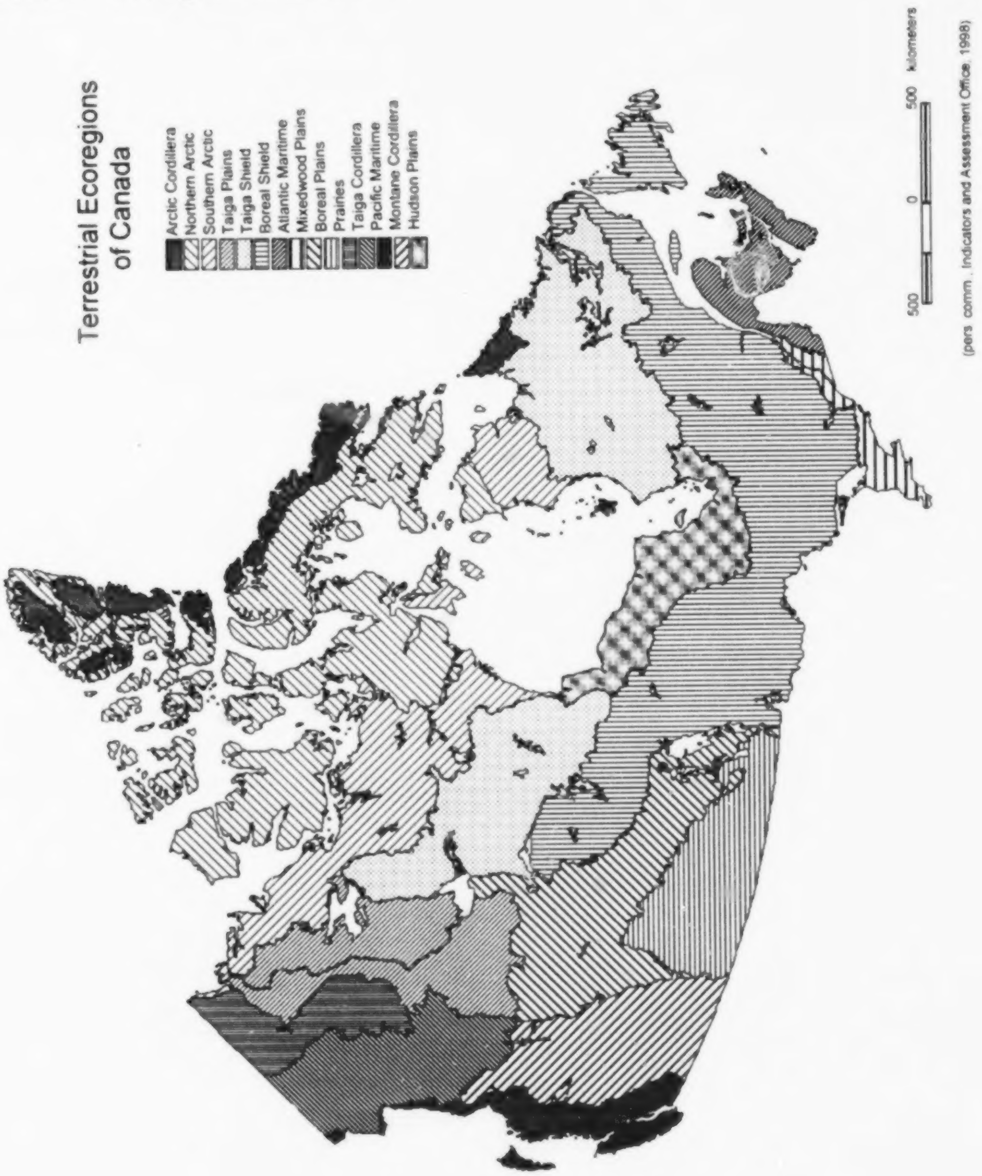

Ecoregion
Drainage Basin
Land-Use
Riverine
Lacustrine
Valley Segment
Stream Reach
Channel Unit
Origin, Location, Linkages
Morphometry
Trophic State
Zone

Figure 3.2 Ecoregions of Canada


Terrestrial Ecoregions of Canada
Arctic Cordillera
Northern Arctic
Southern Arctic
Taiga Plains
Taiga Shield
Boreal Shield
Atlantic Maritime
Mixedwood Plains
Boreal Plains
Prairies
Taiga Cordillera
Pacific Maritime
Montane Cordillera
Hudson Plains
500
0
500 kilometers

(pers comm., Indicators and Assessment Office, 1998)

Maps and descriptions of each ecoregion were produced by Agriculture and Agri-Food Canada and Environment Canada. The ecoregion descriptions were provided through the "National Ecological Framework for Canada" and were written and compiled by the Ecological Stratification Working Group.

3.3.2.2 Drainage Basin and Geographic Scale

Catchments or drainage basins have clear hydrographic boundaries. A drainage basin is defined as that area which has a common outlet for its surface runoff. Although interbasin transfer occurs among biota, the geoclimatic history of large basins (1:2,000,000) are known to create barriers to dispersal through hydrographic divides and climate (Maxwell *et al.* 1995). It is essential to establish the geographic scale appropriate to the study design. In large scale, synoptic surveys in which relationships are sought between landscape features and aquatic biota, the mapping scale for drainage basins is 1:250,000 (Corkum 1989, 1992, 1996, Reynoldson and Rosenberg 1996). These basins are subdivided into progressively smaller sub-basins.

Land/water interactions with respect to sediment and nutrient transport off the land and from upstream sources is integral in developing predictive models that link environmental variables and associated biota. Drainage basins may occur within ecoregions or may cross different ecoregions. Aquatic fauna are more similar to one another in drainages that occupy the same ecoregion than in drainages that occupy different ecoregions (Corkum 1992, Hughes *et al.* 1994).

3.3.2.3 Land-use and Vegetative Buffer Strips

Although ecoregions are defined in terms of climate and natural vegetation, natural vegetation is disturbed with human development. Land-use type is a simple measure of disturbance within the drainage basin. If there is a change in land-use (e.g. land clearing for agricultural uses or logging or fire), the biological assemblages in receiving waters will respond to those changes (Corkum 1992, 1996). Accordingly, site selection should be in drainages with comparable land-use.

The degree (width and type) of a vegetated buffer strip adjacent to rivers and lakes should be recorded at all sampling areas. In reference areas where human disturbance cannot be avoided, the effect of a vegetated buffer strip moderates temperature fluctuations through shading (Budd *et al.* 1987), removes or reduces sediment from runoff (Young *et al.* 1980) and regulates nutrients and metals entering the waterbody (Peterjohn and Correll 1984).

3.3.3 Reference Area Selection in Rivers

The river sampling design provides a framework for characterizing habitats at multiple scales (Meador *et al.* 1993). The framework for rivers is based on how they are organized in hierarchical space and how they change through time (Frissell *et al.* 1986). The riverine system has several hierarchical or nested levels: drainage or catchment basin, valley segment, stream reach and channel unit (Conquest *et al.* 1994) (Figure 3.3).

3.3.3.1 Valley Segment and Stream Order

Valley segments are distinctive sections of drainage basins that possess geomorphic properties and hydrological transport characteristics that distinguish them from other segments (Cupp 1989). Montgomery and Buffington (1993) identified three valley segment types including colluvial (channelized and unchannelized), alluvial and bedrock. Valley segments can be filled with colluvium (sediment and organic matter from landslides) or alluvium (sediment transported by flow). The third valley segment has little soil and is dominated by bedrock.

Valley segments are distinguished by six criteria (Conquest *et al.* 1994):

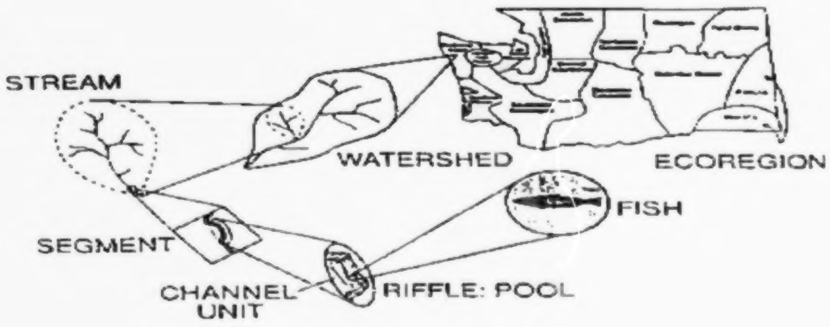
1. Stream order (position in drainage network)
2. Valley slideslope gradient
3. Ratio of valley bottom width to active channel width
4. Channel gradient
5. Stream corridor geomorphic surface deposits
6. Channel pattern.

Channel segments are assigned stream orders (Strahler 1957) for a particular map scale or aerial photograph (e.g., 1:250,000) (Newbury and Gaboury 1993).

3.3.3.2 Stream Reach

Stream reaches consist of homogeneous associations of topographic features and channel geomorphic units (Bisson and Montgomery 1996). Stream reaches can be used to predict local stream response to perturbations (Montgomey and Buffington 1993). Stream reaches are useful in assessing habitat quality, aquatic productivity, fish distributions and stream health (Maxwell *et al.* 1995). Stream reach classification is determined using map scales of 1:12,000 to 1:24,000. Criteria used to classify stream reaches include:

Figure 3.3 Spatially nested hierarchy for stream systems (after Conquest *et al.* 1994)



1. Channel pattern
2. Channel entrenchment
3. Channel width
4. Hydraulic radius
5. Basin area
6. Channel material
7. Stream gradient
8. Bed form
9. Riparian vegetation.

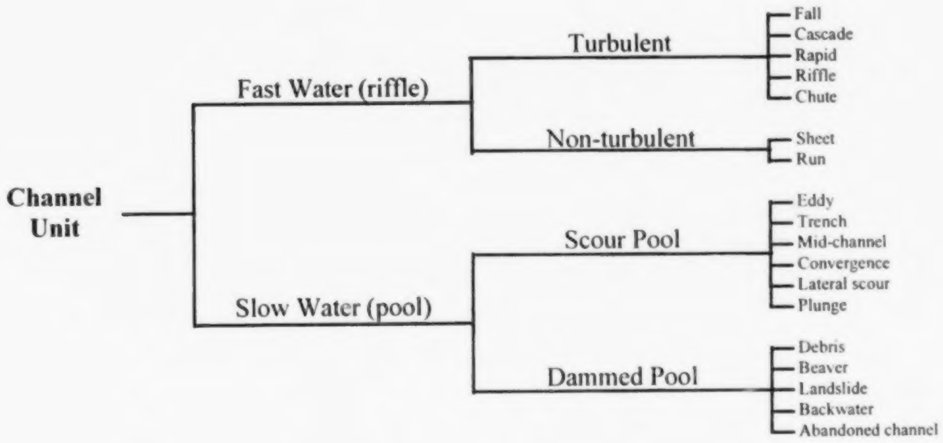
Simpler approaches have been adopted to identify stream reaches. For example, a straight channel has an undulating bed with alternating riffles and pools spaced at repeating intervals of 5-7 channel widths (Leopold *et al.* 1964, Leopold 1994). Newbury (1984) defined a stream reach to be equivalent to six times the channel width.

3.3.3.3 Channel Unit

Channel units are subdivisions of stream reaches that describe uniform microhabitats (depth and flow) and are used to identify factors that limit both invertebrate and fish populations within a stream reach. Hawkins *et al.* (1993) proposed a three-tiered system of channel units in which the first level distinguishes riffles from pools (Figure 3.4). The second level identifies turbulent and non-turbulent riffles and distinguishes between pools formed by scour or dams. Dammed pools retain more sediment and organic debris and have more cover than scour pools. The third subdivision identifies microhabitats based on hydraulic processes and structure. Channel units are typically 10 m or less and typically cannot be mapped at a scale appropriate for land management.

Criteria for subdivision of riffles include:
1. Gradient or water surface profile
2. Percentage of super-critical flow
3. Bed roughness
4. Mean velocity
5. Step development

Figure 3.4 Dendrogram indicating how channel units can be classified with increasing levels of resolution (after Hawkins *et al.* 1993)


Channel Unit
Fast Water (riffle)
Turbulent
Fall
Cascade
Rapid
Riffle
Chute
Non-turbulent
Sheet
Run
Slow Water (pool)
Scour Pool
Eddy
Trench
Mid-channel
Convergence
Lateral scour
Plunge
Dammed Pool
Debris
Beaver
Landslide
Backwater
Abandoned channel

Criteria for subdivision of pools include:
1. Location (main channel or off-channel)
2. Longitudinal and cross-sectional depth profiles
3. Substrate characteristics
4. Pool forming constraints.

3.3.4 Reference Site Selection in Lakes

The geological origin, hydrology, and morphometry (obtained from maps and aerial photographs) of lakes are important in identifying sediment-water interactions and productivity of lakes (Wetzel 1975). Although thermal stratification can be predicted from morphological features, field verification is necessary. The mapping scale for lakes is typically 1:24,000 or 1:63,000 (Maxwell *et al.* 1995).

3.3.4.1 Origins, Location and Hydrological Linkages

Reference and exposure lakes should have the same origin, location and hydrological linkages. Lake geology ultimately affects the physical, chemical and biological characteristics of waterbodies. Hutchinson (1957) identifies 11 types of geomorphic processes (tectonic, volcanic, landslides, glacial activity, solution, fluviatile, wind, shorelines, organic accumulation, anthropogenic and natural dams, meteorite impact). Surface geology and location (altitude, latitude, and longitude) affect lake chemistry and thermal regimes (Winter 1977). These variables, which can be obtained from maps, are used to predict biological composition and productivity of lakes (Dolman 1990, Winter and Woo 1990). Hydrological linkage refers to the "connection of a lake to surface or ground water" and can forecast information about lake biota (Maxwell *et al.* 1995). Maxwell *et al.* (1995) describe three types of hydrological linkages: riverine linkage (outlets and/or inlets or unconnected), groundwater linkage (gaining, losing, neutral or no recharge), and water storage regime (perennial or intermittent).

3.3.4.2 Morphometry

Lake morphometry has been used historically to predict fish yields (Ryder 1965, Kerr and Ryder 1988) and to determine species diversity (Eadie and Keast 1984, Marshall and Ryan 1987). With the exception of depth (and volume), other features can be obtained from maps. Hypsographic (cumulative depth-area or cumulative depth-volume) graphs are useful for comparing basin shapes of lakes and predicting surface area or volume for water level control of reservoirs. Common morphological features of lakes include surface area, volume, mean and maximum depth, shoreline development, and hydraulic residence time.

3.3.4.3 Trophic State

Many lake classification systems are based on a measure of productivity (oligotrophy, mesotrophy, eutrophy). A forth lake type (dystrophy) is used to describe lakes that receive large amounts of organic matter from external sources; these lakes are heavily stained and are known as brown-water lakes. Productivity of dystrophic lakes is low and so some limnologists group dystrophic lakes as a subclass of oligotrophic lakes. The following variables have been used to describe the trophic status of lakes:

- Dissolved oxygen
- Thermal mixing (lake stratification)
- Total phosphorus

- Soluble reactive phosphorus
- Total nitrogen
- Nitrite + nitrate
- ammonium
- Chlorophyll *a*
- Transparency
- Organic matter

3.3.4.4 Zone

Lakes are subdivided into an open water pelagic zone, a shoreline or littoral zone inhabited by autotrophic vegetation and a deeper benthic region free of vegetation, the profundal zone. The substrate characteristics of reference and exposure areas should be similar.

3.4 Reporting of Field Station Positions

The recommended method to describe locations of sampling stations within the exposure and reference areas relative to the outfall location is in global co-ordinates. The latitude and longitude co-ordinates can be obtained using a variety of methods. In some instances, the co-ordinates with stream-wise distances (e.g. river kilometres) may be useful. The recommended positioning accuracy should be determined on a site-specific basis. In some cases, where there are multiple outfalls, industrial sites may decide to collaborate on their studies.

Recommended specifications for positioning and recording sampling stations are: (1) coordinates should be expressed in geographic Universal Transverse Mercator (UTM) or Modified Transverse Mercator (MTM) using the North American Datum (NAD) of 1983 as the geodetic Datum; (2) a local flat plane survey grid should be used only if it has been referenced to a minimum of three geographic UTM or MTM NAD 83 benchmarks; and (3) bathymetry at effluent discharge points located on the Great Lakes should be referenced to depths below the International Great Lakes Datum (IGLD) using the most recent IGLD. Also, federal or provincial water level gauges should be monitored throughout the survey so that water depths can be accurately corrected to IGLD depths.

Additional stations may be included to better represent spatial patterns in a large zone of effluent mixing, for example, at a location with transects (right, mid, left), major habitats (e.g. pool and rifle) and reference/near-field/far-field sampling areas. Recommendations for replicate samples (separately processed) at each station are in sections 5 and 6 of this document.

3.5 Statistical Considerations

This section deals with general statistical issues while specific considerations for the different EEM components can be found in the pertinent sections.

3.5.1 Hypotheses Testing

Statistically, it is necessary to choose between the null and alternate hypotheses (H_o and H_1), as illustrated in Figure 3.5. The probability of rejecting H_o (accepting H_1) when H_o is really true is designated α. This is the probability of reaching a false positive conclusion of impact, and represents risk to the mill. The probability of not rejecting H_o when H_o is really false is designated β. This is the probability of reaching a false negative conclusion of no impact when there is a true impact of magnitude Δ, and represents the risk to the environment. The power ($1-\beta$) of the statistical test, for any Δ, is the probability of detecting a true impact of this magnitude.

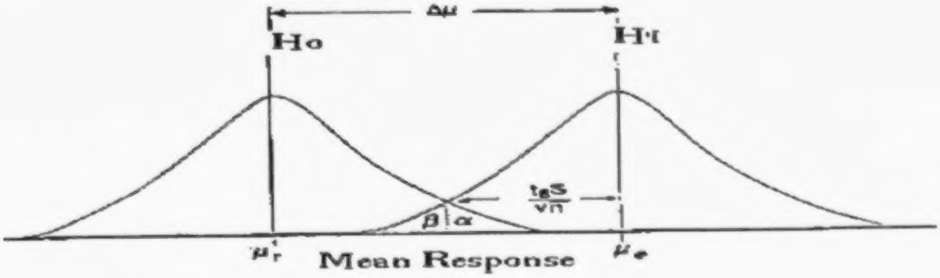
Thus within a study, the null hypothesis about exposure area and reference area means (μ_e and μ_r) can be tested as follows:

$$H_o: \mu_e - \mu_r = 0$$

versus the alternate hypothesis:

$$H_1: \mu_e - \mu_r \neq 0$$

Figure 3.5 Study areas risk (α) and environmental risk (β)



3.5.2 Components of Variation

The standard deviation (SD) term used in all the above calculations of sample size represents the total variation that is expected to be associated with estimates of the area means. In the typical one-stage sampling designs being considered here, it is computed as a pool of SD values within areas, from a set of measurements (one per station) within each area. It includes three hierarchical components, whether or not all components are estimated:

- variation among stations within areas (equals the total of variation if no replicate samples are taken)
- variation among samples (field replicates) within stations
- analytical variation (among laboratory replicates).

If any of these components increases substantially between study design and execution, the desired resolution may not be achieved. Consequently, it is useful to estimate individual components at the design stage and, following execution, to check for changes in the system (Steel and Torrie 1980).

Typically, variation among stations within areas will be the most important component of variation. Variation among samples within stations is generally beyond the control of the sampler and the only way to reduce the contribution of this component to the total variation is by increasing the number of samples per station. This source of variation is less important than the variation among stations (but see Section 6.1 for discussion of sampling effort at a station for ICSs). Analytical variation (variation among laboratory replicates) is often a minor component of the total variation and may be amenable to control. Taylor (1987) recommends inclusion of an allowance for analytical bias in the study design. The square of this allowance can be added to the preliminary estimate of SD^2 to obtain a new SD^2 representing the overall uncertainty that is expected to be associated with measurement of area means. Generally, analytical bias should be small relative to the overall measurement uncertainty (sampling plus analytical plus environmental). The laboratory can suggest a reasonable bias allowance for use in design and as a data quality objective.

3.5.3 Power Analyses and Critical Effect Size Considerations

One of the most critical aspects of Environmental Effects Monitoring (EEM) for pulp and paper mills is the decision on the magnitude of an environmental impact that is deemed to be ecologically important (Mapstone 1996). This critical effect size (ES_{crit}) essentially sets the "bottom-line" in terms of the minimum effect that will trigger management action (Keough and Mapstone 1995). In addition, societal values can play an equally important role in determining ES_{crit} (e.g., economics, aesthetics).

Ideally, a monitoring program should be designed to provide a reasonably high probability of statistically detecting a predetermined ES_{crit} if it has occurred (i.e., the power of the test should be high). The sample size required to achieve this probability level can be determined using power analysis. Two key elements in such analyses are Type I and Type II errors. Type I error occurs (at probability α) if the null hypothesis is rejected when in fact it is true and Type II error occurs (at probability β) if the null hypothesis is accepted when it is false. The power of a statistical test is $1-\beta$, the probability associated with correctly rejecting the null hypothesis when it is false. Two of the relevant power analysis equations and suggested levels for α and β are presented in section 6.1.5.5.

Quantitative guidelines providing ES_{crit} values have not been widely developed for ecological communities, including benthic invertebrate assemblages in either marine or freshwater systems (U.S. Environmental Protection Agency 1990; Elliott 1996). In large part, this has resulted from the complex nature of aquatic communities and the difficulty in making quantitative predictions of the expected magnitude of change following an impact, such as might result from the discharge of pulp and paper mill effluents. Despite these difficulties, some attempt is needed at providing initial estimates of scientifically grounded ES_{crit} values. For further discussions of Power Analysis and its application to the National EEM program in Canada please review Lowell 1997 and section 6 for benthos.

3.6 References

Bisson, P.A. and D.R. Montgomery. 1996. Valley segments, stream reaches, and channel units. Pp. 23-52. In Methods in Stream Ecology (Hauer, F.R. and G.A. Lamberti, eds.). Academic Press, San Diego. 674p.

Budd, W.W., Cohen, P.L., Saunders, P.R. and F.R. Steiner. 1987. Stream corridor management in the Pacific Northwest: 1. Determination of stream-corridor widths. Environmental Management 11:587-597.

Conquest, L.L., Ralph, S.C. and R.J. Naiman. 1994. Implementation of large-scale stream monitoring efforts: sampling design and data analysis issues. Pp. 69-90. In Biological Monitoring of Aquatic Systems (Loeb, S.L. and A. Spacie, eds.). Lewis Publ., Boca Raton. 381p.

Corkum, L.D. 1989. Patterns of benthic invertebrate assemblages in rivers of northwestern North America. Freshwater Biology 21: 191-205.

Corkum, L.D. 1992. Spatial distributional patterns of macroinvertebrates along rivers within and among biomes. Hydrobiologia 239: 101-114.

Corkum, L.D. 1996. Responses of chlorophyll *a*, organic matter and macroinvertebrates to nutrient additions in rivers flowing through agricultural and forested land. Arch. Hydrobiol. 136: 391-411.

Cupp, C.E. 1989. Valley segment type classification for forested lands of Washington. Timber, Fish & Wildlife AM-89-001. 42p.

Dolman, W.B. 1990. Classification of Texas reservoirs in relation to limnology and fish community associations. Trans. Am. Fish. Soc. 119: 511-520.

Eadie, J.M. and A. Keast. 1984. Resource heterogeneity and fish species diversity in lakes. Can. J. Zool. 62:1689-1695.

Elliott, M. 1996. The derivation and value of ecological quality standards and objectives. Marine Pollution Bulletin 32:762-763.

Frissell, C.A., Liss, W.L., Warren, C.E., and M.D. Hurley. 1986. A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. Environmental Management 10:199-214.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley & Sons, New York. 257p.

Hawkins, C.P. Kershner, J.L., Bisson, P.A., Bryant, M.D., Decker, L.M., Gregory, S.V., McCullough, D.A., Overton, C.K., Reeves, G.H., Steedman and M.K. Young. 1993. A hierarchical approach to classifying stream habitat features. Fisheries 18: 3-12.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. Pp. 31-47. In Biological Assessment and Criteria: tools for water resource planning and decision making (Davis, W.S. and T.P. Simon, eds.). Lewis Publishers, Boca Raton. 415p.

Hughes, R.M., Heiskary, S.A., Matthews, W.J. and C.O. Yoder. 1994. Use of ecoregions in biological monitoring. Pp. 125-149. In Biological Monitoring of Aquatic Systems (Loeb, S.L. and A. Spacie, eds.). Lewis Publ., Boca Raton. 381p.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monog. 54:187-211.

Hutchinson, G.E. 1957. A treatise on limnology. Vol I. Geology, physics, and chemistry. John Wiley & Sons Inc. N.Y. 1015p.

Keough, M.H. and B.D. Mapstone. 1995. Protocols for designing marine ecological monitoring programs associated with BEK mills. National Pulp Mills Research Program. Technical Report No. 11.

Kerr, S.R. and R.A. Ryder. 1988. The applicability of fish yield indices in freshwater and marine ecosystems. Limnol. Oceanogr. 33:973-981.

Leopold, L.B. 1994. A view of the river. Harvard University Press, Cambridge. 297p.

Leopold, L.B., Wolman, G.M. and J.P. Miller 1964. Fluvial processes in geomorphology. W.H. Freeman & Co., San Francisco. 522p.

Lowell , R.B. 1997. Discussion paper on critical effect size guidelines for EEM using benthic invertebrate communities. Report to the Environmental Effects Monitoring Program. EEM/1997/9

Mapstone, B.D. 1996. Scalable decision criteria for environmental impact assessment: Effect size, Type I and Type II errors. In Detecting Ecological Impacts: Concepts and Applications in Coastal Habitats (Schmitt, R.J. and C.W. Osenberg, eds.), Academic

Press, San Diego. pp.67-80.

Marshall, T.R. and P.A. Ryan. 1987. Abundance and community attributes of fishes relative to environmental gradients. Can. J. Fish. Aquat. Sci. 44:196-215.

Maxwell, J.R., Edwards, C.J., Jensen, M.E., Paustian, S.J., Parrott, H. and D.M. Hill. 1995. A hierarchical framework of aquatic ecological units in North America (Nearctic Zone). Gen. Tech. Rep. NC-176. St. Paul, MN. U.S. Department of Agriculture, Forest Service, North Central Forest Experimental Station 72p.

Meador, M.R., Hupp, C.R., Cuffney, T.F. and M.E. Gurtz. 1993. Methods for characterizing stream habitat as part of the national water-quality assessment program. U.S. Geological Survey. Open-File Report 93-408, Raleigh, North Carolina.

Montgomery, D.R. & J.M. Buffington. 1993. Channel classification, prediction of channel response, and assessment of channel condition. Washington State Timber/Fish/Wildlife Agreement. Report TFW-SH10-93-002, Department of Natural Resources, Olympia, WA.

Newbury, R.W. 1984. Hydrological determinants of aquatic insect habitats. Pp. 323-357. In The Ecology of Aquatic Insects (Resh, V.H and D.M.Rosenberg, eds.). Praeger, New York.

Newbury, R.W. & M.N. Gaboury. 1993. Stream analysis and fish habitat design. A field manual. Newbury Hydraulics Ltd., The Manitoba Habitat Heritage Corporation, Manitoba Fisheries Branch.

OEPA (Ohio Environmental Protection Agency). 1990. Ohio Water Resource Inventory. Ohio Environmental Protection Agency, Columbus.

Omernik, J.M. 1995. Ecoregions: a spatial framework for environmental management. Pp. 49-62. In Biological Assessment and Criteria. Tools for Water Resource Planning and Decision Making (Davis, W.S. and T.P. Simon, eds.). Lewis Publishers, Boca Raton 415p.

Peterjohn, W.T. and D.L. Correll. 1984. Nutrient dynamics in an agricultural watershed: observation on the role of a riparian forest. Ecology 65: 1466-1475.

Resh, V.H. 1995. Freshwater benthic macroinvertebrates and rapid assessment procedures for water quality monitoring in developing and newly industrialized countries. Pp. 167-177. In Biological Assessment and Criteria. Tools for Water Resource Planning and

Decision Making (Davis, W.S. and T.P. Simon, eds.). Lewis Publishers, Boca Raton 415p.

Reynoldson, T.D. and D.M. Rosenberg. 1996. Sampling strategies and practical considerations in building reference data bases for the prediction of invertebrate community structure. Pp. 1-31. In Study Design and Data Analysis in Benthic Macroinvertebrate Assessments of Freshwater ecosystems Using a Reference Site Approach (Bailey, R.C., Norris, R.H. and T.B. Reynoldson, eds.). Technical Information Workshop North American Benthological Society, 44th Annual Meeting, Kalispell, Montana.

Ryder, R.A. 1965. A method for estimating the potential fish production of north temperate lakes. Trans. Am. Fish. Soc. 94:214-218.

Simon, T.P. 1991. Development of biotic integrity expectations for the ecoregions of Indiana. I. Central Corn Belt Plain. U.S. Environmental Protection Agency, Region V, Environmental Sciences Division, Monitoring and Quality Assurance Branch. Ambient Monitoring Section, Chicago, IL. EPA 905/9-91/025.

Steel, R.G.D. and J.H. Torrie. 1980. Principles and Procedures of Statistics. 2nd ed. McGraw-Hill, Toronto, Ontario.

Strahler, A.N. 1957. Quantitative analysis of watershed geomorphology. Am. Geophys. Union Trans. 38:913-920.

Taylor, J.K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers Inc., Chelsea, Michigan.

Underwood, A.J. 1991. Beyond BACI: Experimental designs for detecting human environmental impacts on temporal variations in natural populations. Australian Journal of Marine and Freshwater Research 42: 569-587.

U.S. EPA (Environmental Protection Agency). 1990. Biological Criteria: National Program Guidance for Surface Waters. Criteria and Standards Division. Office of Water Regulations and Standards, Washington, D.C.

Wetzel, R.G. 1975. Limnology. W.B. Saunders Co., Philadelphia, PA 743p.

Wickware, G.M. & C.D.A. Rubec. 1989. Ecoregions of Ontario. Ecological Land Classification Series, No. 26. Environment Canada. 37p.

Wiken, E. 1986. Terrestrial Ecotones of Canada. Ecological Land Classification Series Number 19, Environment Canada, Ottawa.

Winter, W.T. 1977. Classification of hydrological settings of lakes in the north-central United States. Water Resources Research 134:753-767.

Winter, W.T. and M.K. Woo. 1990. Hydrology or lakes and wetlands. Pp. 159-187. In Surface Water Hydrology (Wolman, M.G. and H.C. Riggs, eds.). The Geological Society of America. Boulder, CO.

Yoder, C.O. 1989. The development and use of biological criteria for Ohio surface waters. Pp. 139-146. In Water Quality Standards for the 21st Century, U.S. Environmental Protection Agency. Office of Water, Washington, D.C. 263p.

Yoder, C.O. 1991. Answering some concerns about biological criteria based on experiences in Ohio. Pp. 95-104. In Water Quality Standards for the 21st Century, U.S. Environmental Protection Agency. Office of Water, Washington, D.C. 251p.

Young, R. A., Huntrods, T. And W. Anderson. 1980. Effectiveness of vegetated buffer strips in controlling pollution from feedlot runoff. J. Environmental Quality 9: 483-497.

4 General Quality Assurance and Quality Control (QA/QC) for Conducting EEM

Quality assurance (QA) encompasses a wide range of internal and external management and technical practices designed to ensure data of known quality commensurate with the intended use of the data.

Quality control (QC) is an internal aspect of quality assurance. It includes the techniques used to measure and assess data quality and the remedial actions to be taken when data quality objectives (DQOs) are not realized. Within the context of a particular study, assurance of adequate data quality is only possible when DQOs have been defined. Users of the data should play a lead role in defining DQOs for a study and in ascertaining whether laboratory quality control limits are consistent with these objectives.

External quality assurance activities include participation in relevant inter-laboratory comparisons and audits by outside agencies. Outside audits may be based on performance in analysis of standard reference materials, or on general review of practices as indicated by documentation of sampling, analytical and QA/QC procedures, test results, and supporting data.

Total Quality Management is one of the aims of QA/QC. This concept stresses the importance of DQOs which encompass all components of uncertainty in data generation, and of data quality check points at all stages from project planning through sampling, analysis, and data interpretation. It requires integration of these activities with real-time quality checks and corrective actions.

The concepts and basic elements of a good measurement system have also been discussed elsewhere (e.g. OECD 1981, 1989; ASTM 1988; Environment Canada 1991). Examples of checklists are provided for evaluation of QA/QC for aquatic biology studies (Table 4.1), aquatic toxicology studies (Table 4.2), and physical and chemical analyses (Table 4.3).

Table 4.1 Example of quality assurance checklist for aquatic biology studies (check or enter footnote number to provide details or to describe deficiencies.)

Quality Management

_______ Standard operating procedures (SOP) available (sampling, processing and QA methods, revisions).
_______ Quality management plan for institution (may be part of SOP).
_______ Responsibilities, authorities, qualifications defined for each position.
_______ Quality assurance officer with adequate authority for corrective action.
_______ Data quality objectives for project (sensitivity, accuracy, precision, total uncertainty).

Quality Assurance Functions Performed

_______ Review of QC sample results (permanent record, replicates, exchanges, verifications).
_______ Performance assessment review (blind samples).
_______ Knowledge testing and training sessions (field and laboratory).
_______ Review of test sample data (transcription and logic checks performed).
_______ Report approval mechanism (signatures).
_______ Quality assurance reports (corrective actions indicated).
_______ Interlaboratory and accreditation studies (record of participation).

Field Operations

_______ Preliminary sampling (return for effort, environmental strata, variability).
_______ Sampling design (based on objectives, adequate replication).
_______ Instrument calibration and maintenance (records, methods available)
_______ Crew training and evaluation (video records and comparisons).
_______ Sampling equipment (appropriate, consistent, cleaning and maintenance records).
_______ Collection, preservation, shipping, storage (methods, adequate labels, custody records).
_______ Additional effort checks (confirm little gain of information with extra effort).

Sample Processing, Quality Control, and Reporting

_______ Sorting and subsampling (methods, adequate records, error estimates).
_______ Sorting verification (confirm few organisms overlooked).
_______ Taxonomy and enumeration (appropriate keys, records, reference collection).
_______ Replicate processing (within and between personnel, error estimates).
_______ Exchange samples (between laboratories, comparability assessed).
_______ Taxonomic verification (by recognized experts).
_______ Fish age verification (by recognized experts).
_______ Archives (for samples, voucher specimens).
_______ Data handling and reporting (entry checks, missing value codes, methods, QC data).

Details/Deficiencies

Footnote #	Description

Table 4.2 Example of quality assurance checklist for aquatic toxicology studies (check or enter footnote number to provide details or to describe deficiencies.)

Quality Management

______ Standard operating procedures (SOP) available (sampling, test and QA methods, revisions).
______ Quality management plan for institution (may be part of SOP).
______ Responsibilities, authorities, qualifications defined for each position.
______ Quality assurance officer with adequate authority for corrective action.
______ Data quality objectives for project (sensitivity, accuracy, precision, total uncertainty).

Quality Assurance Functions Performed

______ Review of QC sample results (permanent record, replicates, reference toxicants).
______ Performance assessment review (blind samples).
______ Knowledge testing and training sessions (field and laboratory).
______ Review of test sample data (transcription and logic checks performed).
______ Report approval mechanism (signatures).
______ Quality assurance reports (corrective actions indicated).
______ Interlaboratory and accreditation studies (record of participation).

Field Operations

______ Preliminary sampling (return for effort, environmental strata, variability).
______ Sampling design (based on objectives, adequate replication).
______ Instrument calibration and maintenance (records, methods available).
______ Crew training and evaluation (video records and comparisons).
______ Sampling equipment (appropriate, consistent, cleaning and maintenance records).
______ Collection, preservation, shipping, storage (methods, adequate labels, custody records).

Laboratory Operations

______ Instrument calibration and maintenance (records, methods available).
______ Animal care and acclimation (records, methods available, tank materials).
______ Reagents and working solutions (labels, acquisition, preparation, expiry dates).
______ Glassware cleaning adequate.
______ Separation of areas (receiving, storage, preparation, testing, laboratory floor plan).
______ Temperature, light, ventilation control (records, adequate devices).
______ Dilution water supply (source, treatment, chemical check records).
______ Food supply (source, chemical check records).
______ Sample information system (log-in, labelling methods, holding times recorded).
______ Safety procedures adequate.

Testing, Quality Control, and Reporting

______ Test record sheets (concentrations, daily responses, test conditions).
______ Dose selection guidelines (range tests, concentration intervals).
______ Dose verification (frequency, methods for mixtures).
______ Dilution water controls (use of data, test acceptance criteria).
______ Test replication (frequency, range control limits).
______ Reference toxicant tests (type, frequency, control limits).
______ Warning charts or tables (limit determination, out-of-control actions).
______ Method selection, guidelines for endpoint calculations.
______ Data, handling and reporting (confidence limits, methods, test conditions, QC data).

Details/Deficiencies

Footnote #	Description

Table 4.3 Example of quality assurance checklist for physical and chemical analyses (check or enter footnote number to provide details or to describe deficiencies.)

Quality Management

______ Standard operating procedures (SOP) available (sampling, analysis and QA methods, revisions).
______ Quality management plan for institution (may be part of SOP).
______ Responsibilities, authorities, qualifications defined for each position.
______ Quality assurance officer with adequate authority for corrective action.
______ Data quality objectives for project (sensitivity, accuracy, precision, total uncertainty).

Quality Assurance Functions Performed

______ Review of QC sample results (permanent record, replicates, blanks, standards).
______ Performance assessment review (blind samples).
______ Knowledge testing and training sessions (field and laboratory).
______ Review of test sample data (transcription and logic checks performed).
______ Report approval mechanism (signatures).
______ Quality assurance reports (corrective actions indicated).
______ Interlaboratory and accreditation studies (record of participation).

Field Operations

______ Preliminary sampling (return for effort, environmental strata, variability).
______ Sampling design (based on objectives, adequate replication).
______ Instrument calibration and maintenance (records, methods available).
______ Crew training and evaluation (video records and comparisons).
______ Sampling equipment (appropriate, consistent, cleaning and maintenance records).
______ Collection, preservation, shipping, storage (methods, adequate labels, custody records).

Laboratory Operations

______ Instrument calibration and maintenance (records, methods available).
______ Reagents and working solutions (labels, acquisition, preparation, expiry dates).
______ Glassware cleaning adequate.
______ Separation of areas (receiving, storage, preparation, analysis, laboratory floor plan).
______ Temperature, light, ventilation control (records, adequate devices).
______ Sample information system (log-in, labelling methods, holding times recorded).
______ Safety procedures adequate.

Analysis, Quality Control, and Reporting

______ Bench record sheets (dilutions, pre-concentrations, readings, calculations).
______ Trip and laboratory blanks (use of data, acceptance criteria).
______ Field and laboratory replicates (frequency, range control limits).
______ Standard reference materials (types, frequency, control limits).
______ Control charts or tables (limit determination, out-of-control actions).
______ Blank and matrix spikes (types, frequency, use of data).
______ Data handling and reporting (blank, recovery correction, significant figures, flags, QC data).

Details/Deficiencies

Footnote # Description

4.1 Quality Management

A clearly defined organizational structure is an essential part of a QA/QC program. Specific individuals in the organization are assigned QA/QC functions, data reporting channels are established, and the authority for specific corrective actions to be taken when established performance limits or data quality objectives are not satisfied is identified.

The suggested method to describe this structure is a quality management plan situated in a quality assurance plan (QAMS 1980). Quality management plans normally include a QA policy statement signed by a person with budgetary authority.

Quality management plans may be developed for single institutions, laboratories (e.g. OMOE 1985) or for inter-agency programs (e.g. LRTAP 1986; IJC 1988). They are particularly critical in an inter-agency context. Similarly, quality assurance plans (QAPs) may be developed for an institution or a project.

4.1.1 Standard Operating Procedures

Standard Operating Procedures (SOPs) are written detailed methods to be followed for sample collection or analysis. The SOP document usually contains more technical detail on specific methods than the QAPs and cites published methods manuals as appropriate. SOPs can be referenced in any QAP.

Each SOP should be a written method accessible to each analyst. SOPs should be based on procedures developed by a standard-setting organization such as Environment Canada, the U.S. EPA, the American Society for Testing and Material (ASTM), and American Public Health Association (APHA). Where methods are not well validated, it is recommended that the SOP be thoroughly referenced to the relevant literature and contain all the elements outlined in Canadian Association for Environmental Analytical Laboratories (CAEAL) Inc. (1991) where applicable. In-house validation data should be appended to the SOP and should contain the quality assurance and quality control procedures. These include the types and frequencies of quality control samples to be analyzed, expected levels of precision, accuracy, recovery, and method detection limits.

While chemical analysis procedures tend to be reasonably well-documented, sampling procedures in general, and sampling design in particular, are often overlooked. Sampling error is usually a large component, and often the dominant component of uncertainty in environmental measurements. Standard operating procedures that include field operations will help to reduce this uncertainty or at least ensure that it is quantified.

4.1.2 Data Quality Objectives

Data quality objectives (DQOs) are normally derived from intended data uses (e.g. hypotheses to be tested, summary statistics involved, and total uncertainty that can be tolerated). Total uncertainty includes imprecision (sampling, analytical, environmental) plus any analytical bias that may occur (Taylor 1987). Objectives can be established for each component, as well as for total uncertainty, and should be incorporated into the QA project plans. The various components of imprecision can be estimated using field replicate and laboratory replicate data.

4.2 Quality Assurance Functions

Quality assurance functions, personnel responsible for each QA function, and corrective actions when performance limits are exceeded should be identified in the quality management plan.

Quality control activities define the boundaries of acceptable performance for the measurement system, and include the routine checks (data quality measures) that indicate whether the system is performing to specification. Data reporting generally stops and corrective actions are initiated when the system goes out of control. Range and average control charting methods have been described elsewhere (OMOE 1984; ASTM 1985, 1986; Dux 1986).

Data quality measures should be defined in the same terms as data quality objectives (DQOs), so that the two can be compared in project evaluation (QAMS 1986).

The principles of quality management and quality assurance are applicable to all measurement systems for which sampling and analytical performance checkpoints are defined. However, some sources of error are unique to ecological, toxicity, physical/chemical or bacteriological measurement systems, and may require unique data quality checks.

4.3 Field Operations

Field methods should specify sampling equipment and protocols appropriate to the study. Consistency in sampling equipment reduces biases in terms of collection efficiency and type of organism captured. Methods should also specify proper use of equipment, location of stations, sample containers, labelling and preservation of samples, and any constraints on sample storage (e.g. Tetra Tech, Inc. 1987). Finally, frequencies of field replicate samples should be specified.

Contamination during chemical sampling can be checked with trip blanks and equipment rinsates. Trip blanks are prepared in the laboratory, taken to the field, collected as regular samples and returned to the laboratory for analysis. Rinsates are water or solvent samples processed through the sample collection device(s) in the field just prior to sampling. Such samples provide a "background" chemical concentration. Contamination problems can be minimized by appropriate washing and checking procedures for sample containers and sampling devices, or sterilization procedures for microbiological sampling equipment.

Additional effort checks can be performed in the field during sampling, and serve to confirm the return-for-effort anticipated on the basis of preliminary field work. For example, if it is estimated that three grabs per station are sufficient to achieve a data quality objective to retrieve 95% of benthic species present at any station, more grabs can be collected at a few selected stations and analyzed. Determination of a plateau from these samples determines whether the first three grabs were sufficient to achieve the 95% objective. Stevenson and Lowe (1986) provide an example related to periphyton sampling.

4.4 References

ASTM. 1985. Standards on Precision and Bias for Various Applications. 2nd ed. American Society for Testing and Materials, Philadelphia.

ASTM. 1986. Manual on Presentation of Data and Control Chart Analysis. Committee E-11 on Statistical Methods. ASTM Special Technical Publication 15D. American Society for Testing and Materials, Philadelphia.

ASTM. 1988. Draft Standard Practice for Generation of Environmental Data. Committee E-47 on Biological Effects and Environmental Fate. Subcommittee E47.05 on Quality Assurance. American Society for Testing and Materials, Philadelphia.

CAEAL (Canadian Association of Environmental Analytical Laboratories). 1991. Code of Practice and QA Manual for Laboratory Analysis of Sewage Treatment Effluent in Support of the MISA Program; Draft report prepared for CAEAL and the Ontario Ministry of the Environment by Zenon Environmental Laboratories.

Dux, J.P. 1986. Handbook of Quality Assurance for the Analytical Chemistry Laboratory. Van Nostrand Reinhold Co., New York.

Environment Canada. 1991. Quality Assurance Guidelines for Biology in Aquatic Environmental Protection. Prepared by Dr. D. Hart, Beak Consultants Limited for Dr. A.S.Y. Chau, National Water Research Institute, Burlington, Ontario.

IJC (International Joint Commission). 1988. Quality Management Plan for Great Lakes International Surveillance with Implementation Guidelines. May 1988. (draft).

LRTAP (Long-Range Transport of Air Pollutants). 1986. Long-Range Transport of Air Pollutants. Quality Management Plan. Federal/Provincial Research and Monitoring Coordinating Committee.

OECD (Organisation for Economic Cooperation and Development). 1981. Decision of the Council Concerning Mutual Acceptance of Data in the Assessment of Chemicals. Annex 2. OECD Principles of Good Laboratory Practice. In: Guidelines for Testing of Chemicals. Organisation for Economic Cooperation and Development, Paris.

OECD. 1989. Draft Council Decision-Recommendation on Compliance with Principles of Good Laboratory Practice (restricted). Organisation for Economic Cooperation and Development, Paris.

OMOE (Ontario Ministry of the Environment). 1984. Principles of Control Charting. D.E. King and G.C. Ronan, Laboratory Services Branch, Data Quality Report Series. Ontario Ministry of the Environment, Rexdale, Ontario.

OMOE. 1985. Laboratory Services Branch, Quality Management Plan. Ontario Ministry of the Environment, Rexdale, Ontario.

QAMS (Quality Assurance Management Staff). 1980. Interim Guidelines and Specifications for Preparing Quality Assurance Project Plans. Quality Assurance Management Staff, Environmental Protection Agency, Washington, D.C. EPA-600/4-83-084. PB-83-170514.

QAMS. 1986. Development of Data Quality Objectives. Description of Stages I and II. Environmental Protection Agency, Washington, D.C. (Draft).

Stevenson, R.J. and R.L. Lowe. 1986. Sampling and interpretation of algal patterns for water quality assessments. In: Rationale for Sampling and Interpretation of Ecological Data in the Assessment of Freshwater Ecosystems, Isom, B.G. (ed.), STP 894. American Society for Testing and Materials, Philadelphia,

Taylor, J.K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers Inc., Chelsea, Michigan.

Tetra Tech, Inc. 1987. Recommended Protocols for Sampling and Analyzing Subtidal Benthic Macroinvertebrate Assemblages in Puget Sound. A report prepared for U.S. EPA, Region 10, Seattle. Final Report TC-3991-04.

5 Effects on Fish and Fisheries Resources: Fish Survey, Tissue Analyses, and Tainting Evaluation

5.1 Introduction to Fish Survey

A description of methods for the Fish Survey is provided in this section, which includes detailed guidance on:

- the general approach of the fish survey
- survey and sampling design
- survey measurements
- data summary and power analysis
- data quality assurance/quality control

5.1.1 Fish Survey

5.1.1.1 Approach

The purpose of the fish survey component of EEM for cycles subsequent to Cycle One will be to assess the adequacy of the regulations by examining the effects of effluents on the survival, growth and reproduction of fish, where "fish" are defined as finfish, crustaceans and bivalves. The Fish Survey (FS) focuses on monitoring whole-organism parameters for two sentinel fish species resident in the receiving environment and in the reference area. The status of a sentinel species, as indicated by a variety of characteristics measured on individual fish, should reflect the overall condition of the aquatic environment in which the fish reside. The FS will, over the course of the EEM iterative process, identify whether there are differences in the life-history parameters of fish exposed to mill effluent. Fish survey measurements are presented in Table 5.1.

There are several advantages to this approach, including that it:

- uses traditional fisheries information (individual fish measurements)
- focuses on a few select fish species and does not depend on intensive (and costly) sampling of fish communities or populations
- is cost effective

Table 5.1 Fish survey measurements

Measurement Requirement	Expected Precision	Reporting Requirement
Length (fork or total or standard)*	+/- 0.2cm	Individual measurements, mean, standard deviation
Total body weight (fresh)	+/- 5.0%	Individual measurements, mean, standard deviation
Age	+/- 1 year (10% to be independently confirmed)	Individual measurements, mean, standard deviation
Gonad weight	+/- 1.0%	Individual measurements, mean, standard deviation
Egg size	+/- 1.0%	Weight, minimum sub-sample sizes of 100 eggs
Fecundity**	+/- 1.0%	Total number of eggs per female
Weight of liver or hepatopancreas	+/- 1.0%	Individual measurements, mean, standard deviation
External condition	NA	Obvious abnormalities, prevalence of lesions, tumors, parasites, etc.
Sex	NA	

* If caudal fin forked, use fork length. Otherwise, use total length. Cases where fin erosion is prevalent, standard length should be used.

** Fecundity can be calculated by dividing total ovary weight by weight of individual eggs (individual egg weight can be estimated by counting the number of eggs in a sub-sample. The sub-sample should contain at least 100 eggs).

5.1.1.2 Previous Cycles

The EEM program is a sequential series of monitoring and interpretation cycles, wherein the requirements of each cycle are dependent on the findings of the previous cycle. The focus of the first cycle of the FS was to conduct a preliminary survey which would allow the investigator to become more familiar with the receiving environment, study area and selected sentinel species and to provide data describing the variability in fish parameters. Thus the results of previous cycles should be used to guide the development of subsequent cycles. No individual EEM cycle alone will provide the necessary information to evaluate the effectiveness of the current regulations. It is a series of cycles that will permit this evaluation on the basis of a scientifically-defensible data base.

The primary objective of the first cycle FS was to obtain an estimate of the variability in fish parameters. This estimate of variability is to be used in power calculations to refine the study designs of subsequent cycles. The results of previous cycles have been used to assess:

- the suitability and capture success of selected sentinel species
- the adequacy of the reference area
- the appropriateness of sampling methods and required equipment.

Interested persons are referred to the Fish Survey Expert Working Group Final Report (Environment Canada 1997c) for more information on Cycle One data.

An effective way of utilizing previous cycles' data in the development of the current cycle study design is through the Fish Survey Decision Tree (Figure 5.1) and associated tables (Tables 5.2 to 5.8). This tree should help determine the success of previous cycles and aid in decision-making for the current cycle. Each mill's EEM representatives or consultants must meet with the regional EEM authority to review the results of the previous cycle's site selection, species selection, fishing effort, etc. and to select the most appropriate options for the current cycle. Variability estimates from Cycle One should be used to calculate the sample size required to detect a difference between exposed and reference fish. A ±25% difference in gonad weight relative to body weight or length (=relative gonad weight) between reference and exposed areas should be adopted as a target effect size, using a power=0.80.

Excessive precision should be avoided when specifying target effect size (ES) and calculating sample sizes for Cycle 2 and subsequent cycles (Section 5.1.6.2). The 25% difference chosen for the target ES should be regarded as approximate; a range (i.e. 20-30%) would be more appropriate. The lower limits for the target ES were set by two considerations:

1) Many fish surveys would not have sufficient statistical power to detect smaller effects (e.g. <20% and certainly <10%), except over many cycles and/or with larger sample sizes than used thus far. As target ES become smaller, and required sample sizes larger, the effects of sampling mortality on the population will eventually exceed any population-level effects from reproductive effects.

2) Differences in relative gonad weight of <20% may occur naturally between unexposed and relatively similar areas. If the target ES is too small, then the risks of false positives (detecting a natural difference between areas and attributing it to the mill effluent) increase.

Figure 5.1 Fish survey decision tree

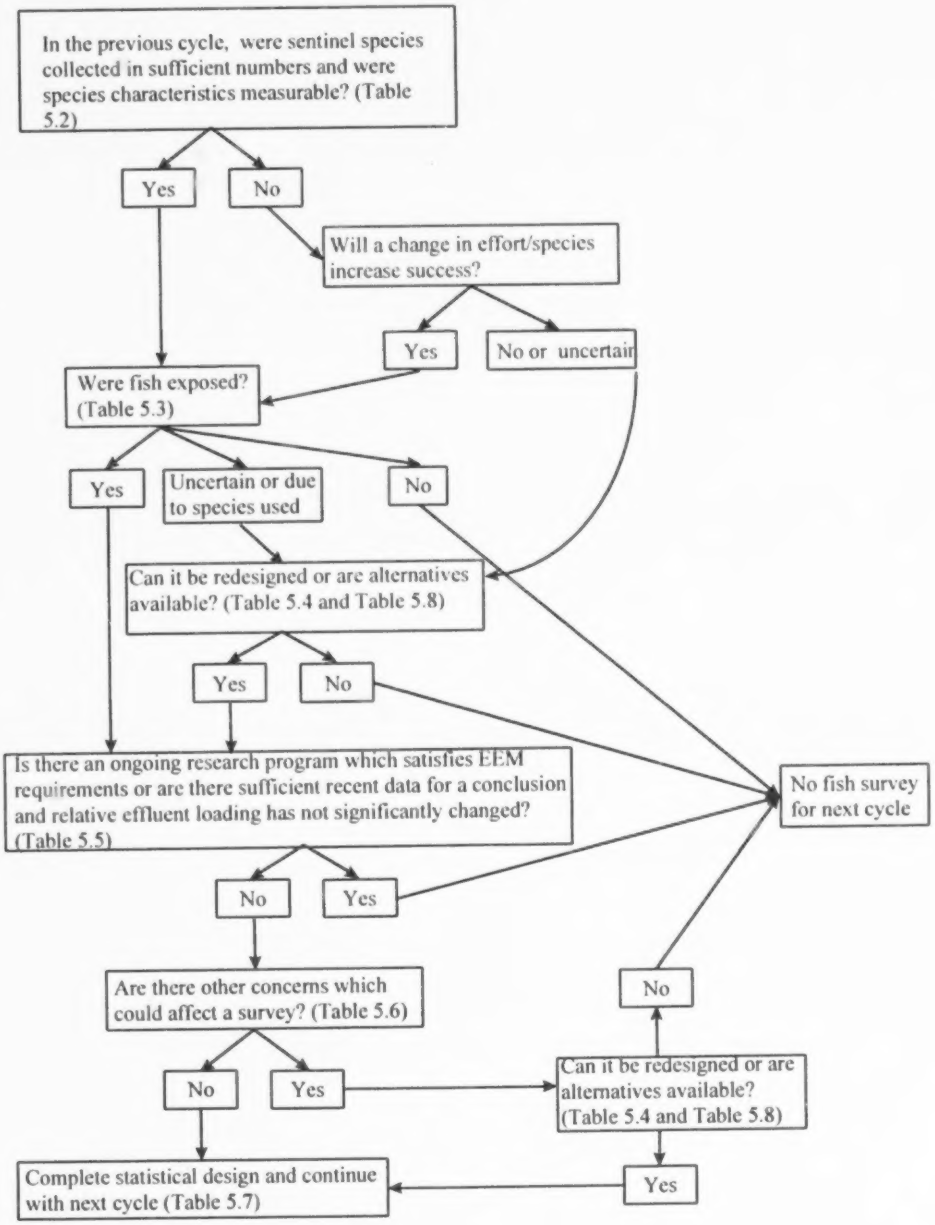

In the previous cycle, were sentinel species collected in sufficient numbers and were species characteristics measurable? (Table 5.2)
Yes
No
Will a change in effort/species increase success?
Yes
No or uncertai
Were fish exposed? (Table 5.3)
Yes
Uncertain or due to species used
No
Can it be redesigned or are alternatives available? (Table 5.4 and Table 5.8)
Yes
No
Is there an ongoing research program which satisfies EEM requirements or are there sufficient recent data for a conclusion and relative effluent loading has not significantly changed? (Table 5.5)
No
Yes
No fish survey for next cycle
Are there other concerns which could affect a survey? (Table 5.6)
No
Yes
Can it be redesigned or are alternatives available? (Table 5.4 and Table 5.8)
No
Yes
Complete statistical design and continue with next cycle (Table 5.7)

Table 5.2 Sentinel species characteristics and criteria for success

Criteria for defining "sufficient numbers" and "suitable characteristics" of sentinel fish species

1. were 20 fish of each sex of 2 species caught in each area? If yes, go to 2
2. was it possible to measure all fish survey parameters on the fish? If yes, go to 3
    - age
    - body weight and length
    - liver or hepatopancreas weight
    - gonad weight
    - egg size
3. were gonad weights of at least 80% of the female fish greater than 1% of the carcass weight? If yes, go to 4
4. if the sentinel species is a fractional spawner are you confident that sampling was not during the spawning period?

Some sentinel species may not meet all these criteria

Table 5.3 Were fish exposed?

If fish were captured in the effluent field, then exposure is obvious, although the period of residency is unknown. In situations where large, statistically significant differences are seen between exposure and reference areas, it is obvious that there has been a significant period of residency and an absence of mixing of fish between areas. Although these differences may not be the result of exposure to mill effluents, they do indicate some significant period of residency in the effluent field, and an absence of mixing between reference and exposure fish. The main concern relates to establishing the period of residency in the area of the effluent field when no obvious barriers are present and there are no differences between areas. Most chemical and biochemical tracers only provide evidence of short-term residency (days to weeks) and they were not used very successfully in the first cycle. The real value of chemical tracers is to establish that reference fish did not spend significant time in the effluent field. If there is doubt about the exposure, then redesign of the survey should be considered to utilize fish which are more likely to remain longer in the effluent field.

Surveys should not be conducted in the case of high dilution:

1. In rivers, - if effluent concentration, based on relative flows, is <0.1% after complete mixing at low flow, and if the effluent is diluted to <1% within 250 m of the discharge.
2. in marine and lake systems, if the effluent is diluted to <1% within 250 m of the discharge.

Table 5.4 Guidance for re-design

1. If fish exposure was uncertain
   a) are less mobile fish available, which would increase likelihood of exposure? If no, go to 1.b).
   b) if no other species are available, would changing the time of sampling increase the likelihood of exposure, using the same species as in Cycle 1? If no, go to 2.

2. Are there other factors (e.g. type of sampling gear) that would increase success? Note: the use of one species may be acceptable. If no, go to 3.

3. Sentinel fish availability is a problem. If other alternatives (See Table 5.8) are not appropriate, then go to 4.

4. Survey can not be redesigned, thus, no fish survey.

Table 5.5 Are there sufficient recent data for a conclusion?

The purpose of the EEM program is to evaluate the adequacy of the Regulations. There are three possible conclusions which can be reached: there is no effect, there is a mill-related effect, or there are confounding factors which prevent determination of the cause of an effect.

1. For determining that there are or are not changes, a minimum of two cycles, with a power and effect size that satisfies the requirements of Table 5.7, are required. Supplementary data could be used to replace one cycle of EEM data if it provides the required information on age, growth and reproduction, and has been collected since the last major process change.

2. Confounding factors - If there is sufficient data to demonstrate that other discharges or contaminant sources are primarily responsible for observed changes or absence of observed changes, and the survey cannot be re-designed to isolate mill effects, then there can be a conclusion that confounding factors prevent interpretation for the purposes of assessing the adequacy of the Regulations.

Table 5.6 Other concerns which could affect a survey

Other discharges are present and survey cannot be re-designed to isolate impact. Some examples include:

1. mill effluent is a minor component (*e.g.* less than 20%) of a combined discharge
2. the mill discharge is located within the effluent field of an adjacent discharger and/or adjacent discharge is located within the mill effluent field, and the mill discharge represents a minor component (*e.g.* less than 20%) of the discharges into the effluent field, and the survey cannot be re-designed to isolate the mill effect.

Other confounding factors may exist which have not been identified.

Habitat unsuitable or unsafe for sampling: *e.g.* extreme tidal fluctuations, high current velocity, etc.
No suitable reference site: *e.g.* due to habitat, fish movements (reference fish exposed).
Permit restrictions or concerns about protected, rare, or endangered species.
Restrictions on sampling such as timing or location preventing sampling of appropriate species or correct timing.

Table 5.7 Power analysis

The sample sizes should be determined using the variability estimates given in Cycle 1, with a target power of 0.8.The study design should be based on the ability to detect a difference of 20-30% in relative gonad size between reference and exposed areas. Other variables to be measured (Table 5.1) will require further monitoring in order to establish effect sizes. If variability estimates are not available, a sample size of 20 males and 20 females from each area would be conservative. If effects were observed in Cycle 1 or subsequent cycles, monitoring must be expanded to delineate the spatial extent of the effects.

Table 5.8 Possible alternatives for the fish survey

No specific alternatives are recommended at this time. However, various alternatives are under review and this Table will be updated as alternatives are considered.

The issues of species selection, site selection, fishing methods and fishing effort are discussed in detail in the following sections.

5.1.2 Selection of Sentinel Monitoring Species

Annex 1 states that a minimum of two sentinel species should be monitored to fulfill the requirements of the Fish Survey (1997b). In addition, at least one of the species selected for the FS should be a benthivore.

The success of the FS depends heavily on the selection of appropriate sentinel fish species. The decision of which species are the most appropriate as sentinel species will be site-specific. A key consideration when selecting a sentinel species is the mobility and residency time of that species (as this determines effluent exposure). Species that are resident in the system for most or all of their life cycle and exhibit territorial behavior or limited mobility relative to the size of the study area are preferred because the observed responses of these species reflect their localized environment. Species that are migratory or spend only a small proportion of their life cycle in the system under investigation (e.g. anadromous salmonids, some marine fishes) are not suitable as sentinel species because exposure to effluent is minimal or transient. This is also true for species that are highly mobile and are likely to be moving in and out of the effluent exposure area. In some cases, it may be possible to select more mobile species (e.g. mountain whitefish - Swanson 1993) as sentinels due to physical constraints which limit movement (e.g. dams, natural barriers, changes in habitat) . In general, the greater the likelihood that a fish species is exposed to effluent, the greater its value as a sentinel species.

In most cases, there are limited choices of potential sentinel species for monitoring and site-specific requirements will dictate the selection. It will often be necessary to obtain the advice of an experienced fisheries biologist with knowledge of fish species present in the study area. The mills and their consultants are encouraged to contact regional federal and provincial government agencies for fisheries information and additional guidance. Very large species should be avoided, since adult abundance of small species will respond more rapidly to stressors that affect fecundity and/or larval survival (Shuter 1990). Species that are harvested for sport or commerce are not suitable as sentinel species (exploitation on a population may obscure effects related to environmental stress).

Utilization of small bodied, resident fish species should be considered in the absence of large bodied fish. Smaller species usually have a more limited home range than large bodied species, and have an increased likelihood of exposure (Gibbons et al. 1998a). However, there are several key considerations which should be kept in mind when designing the field studies using the smaller species (Gibbons et al. 1998a,b).

Species identification of small bodied fish should be verified, especially for cyprinids which can appear very similar without careful examination (Scott 1967; Scott and Crossman 1973; Roberts 1988; Nelson and Paetz 1992; Jenkins and Burkhead 1993; Coad 1995). The smaller organ size of these fish requires a more sensitive balance. Dissecting microscopes may be necessary for removing the organs properly and avoiding extraneous tissue or moisture which could affect results. Dissection is usually easier on live, fresh fish, than frozen specimens. Differentiation between tissues is easier when the tissue is fresh, and separation of the liver and gonads from intestinal tissue is less time consuming.

Small bodied fish usually have a shorter life span. There is generally greater success with traditional age structures such as scales (Table 5.9), but additional structures (such as opercula and otoliths) should be preserved in case the scales should prove inadequate or verification is necessary.

Measurement of fecundity and egg size requires special consideration. Many small species have few, large eggs. Gonadal estimates will be much easier closer to spawning. The timing of sampling will also be affected by residency, and the two factors have to be optimized. The entire gonad should be preserved and fecundity counts conducted with aid of dissecting microscope. Basic information on gonadal development should be available for many species through general references (Scott 1967; Scott and Crossman 1973; Fritz et al. 1975; Roberts 1988; Nelson and Paetz 1992; Jenkins and Burkhead 1993; Coad 1995; Leblanc and Couillard 1995; Leblanc et al. 1998).

Consideration has to be given to reproductive life histories when selecting smaller fish species. There is the same variety of reproductive strategies in small fish as evident in larger fish species. Multiple spawners require special attention because they may spawn more than one clutch of ova per year. The ovary will generally contain two or more class sizes of ova and the spawning season may last from several weeks to more than a month. In some species, the entire ovary mass can be turned over more than once during the spawning season. Incorrect timing of sampling may lead to problems estimating reproductive effort and fecundity in multiple spawners. At the present time, it is assumed that sampling prior to the initiation of the first spawning period for that annual cycle will provide an estimate of reproductive investment relative to the entire year.

Live-bearers require special attention regarding measurement of reproductive variables, but success has been seen with detecting responses using live-bearing species in exposures to Swedish pulp mills (Vetemaa et al. 1997). The estimates of fecundity requires that the gonad be preserved and the number of live and dead embryos be counted. Proper sampling requires some preliminary data on spawning time and gonadal development so that sampling procedures can be optimized. This data may not be

Table 5.9 Anatomical structures commonly used for age determination

Family	Common Name	Structure
Acipenseridae	Sturgeons	Otoliths, pectoral fin ray
Anguillidae	Freshwater eels	Otolith
Atherinidae	Silversides	Otolith
Batrachoididae	Toadfishes	Otolith
Bothidae	Lefteye flounders	Otolith, scale
Caranigidae	Jacks	Otolith
Catostomidae	Suckers	Fin ray, opercular bone
Centrarchidae	Sunfishes	Scale
	Bass	otoliths, 1st three dorsal spines
Cichlidae	Cichlids	Scale
Clupeidae	Herrings	Otolith
Coregoninae	Whitefish	Otoliths, first four marginal pectoral fin rays, scales
Cyprinidae	Minnows and carps	Otoloith, scale
Cyprinodontidae	Killifishes	Scale
Esocidae	Pikes	Scale, cleithra
Gadidae	Codfishes	Otolith
Haemulidae	Grunts	Otolith
Ictaluridae	Bullhead catfishes	Pectoral spine
Lophiidae	Goosefishes	Vertebrae, ray
Malacanthidae	Tilefishes	Otolith
Mugilidae	Mullets	Scale
Percichthyidae	Temperate basses	Scale
Percidae	Walleye, perch	Otoliths, first three dorsal spines, opercular bone, scale
Pleuronectidae	Righteye flounders	Otolith, scale
Rajidae	Skates	Vertebral centrum
Salmonidae	Lake trout	Scale, otolith, subopercula, 1st four marginal pectoral fin rays
Sciaenidae	Drums	Scale
Scombridae	Mackerel	Otolith, fin ray
Serranidae	Sea basses	Scale
Sparidae	Porgies	Scale
Squalidae	Dogfish sharks	Dorsal spine

available for Cycle 2 for live-bearing fish, but should be evaluated on a case-by-case basis.

The *Fisheries Act* also includes shellfish under the definition of fish, and they were used by some mills during Cycle 1. Crabs and lobsters are not appropriate sentinel species because they cannot be reliably aged (Environment Canada 1997b). The suitability of wild mussels and oysters are being evaluated as possible alternatives.

In a limited number of cases, there have been situations where no adult fish can be collected in a receiving environment. If no large bodied fish could be collected, and small-bodied fish were included in attempts, the suitability of juvenile fish may be considered. The use of juvenile fish is being evaluated as a possible alternative. If reproductively-active fish do not reside in an effluent field, the only relevant measures of fish performance are those related to growth and survival of juvenile, resident, exposed fish. The priority considerations are residency and exposure of the juvenile fish.

5.1.3 Sampling Design

5.1.3.1 Minimum Design

Annex 1 (Environment Canada 1997b) specifies the minimum design for the first cycle of the Fish Survey as:

- sampling a minimum of one reference area and one near-field area, and
- collection of a minimum of 20 adult females and 20 adult males of each of two sentinel species per sampling area.

Sample sizes will be set based on the target effect size for reproduction, and effect sizes which could be detected for other parameters calculated and reported. These detectable effect sizes for other parameters should still be compared to natural differences observed in other studies as part of interpretation and study design for subsequent cycles. When information on variability is not available from a previous cycle, the mill should set 20-30 fish as a target sample size and attempt to develop information on variability.

If only one sentinel species is available, approval must be obtained from the regional officer to carry out the study with a single sentinel species.

The following discussion briefly summarizes the development of the minimum design.

Sampling Areas

The minimum number of sampling areas was defined in response to the need to designate a minimum acceptable level of sampling effort that would achieve the objectives of the first cycle and was applicable to all mills. Specifically, it was recognized that sampling reference and near-field areas at some mill sites would be limited due to restrictions posed by heterogenous or dissimilar habitat, access, or other factors which limit sampling opportunities. However, it was unacceptable not to sample at least one reference and one near-field area. Therefore, on a national basis, the minimum design stipulated sampling at least one reference area and at least one near-field area.

For clarity, consideration of the comparison of two zones (effluent and reference) with the possibility of sampling several areas within each zone and sub-sampling units (fish) within each area. If we wished to conduct a true experiment to investigate the difference between effluent and reference we would start with some collection of areas, randomly divide these into two groups and randomly assign the "treatments" to these groups. Obviously this is not possible in an EEM study. The effluent zone is defined by a particular mill discharging into a particular body of water. At this level there is no element of randomization. The definition of the reference zone is so restrictive that we should consider it also as a unique entity. In analysis of variance terms, effluent and reference zones should be considered a fixed effect.

The simplest way to approach the sampling design of the FS is to sample fish from a single reference and near-field zone (i.e. the minimum design). This approach assumes that the habitat within each zone is fairly homogenous such that there is a widespread distribution of fish without recognizable sub-populations. In this case, sampling any location (area) within the zone provides an equal chance of capturing fish (i.e. it is possible to obtain a random sample of fish), and the appropriate error term used to test for differences in areas (future FS cycles) is the "among-fish within-area" variance. Consequently, when conducting the first cycle of the FS, the appropriate level of variability in fish parameters that should be estimated is the variability "among fish".

It is recognized that sometimes the habitat within a zone is not homogenous, and the distribution of the fish population reflects the patchy distribution of suitable habitat. In this case, there is probably good reason to believe that within the zone, the fish population is subdivided into a number of spatially discrete sub-populations which could be considered discrete "areas". However, for the purpose of meeting the objectives of the first cycle, there is little benefit to sampling more than one of the sub-populations or areas. Regardless of whether single or multiple areas within a zone are sampled, the appropriate error term used in statistical comparisons is still the "among-fish within-area" variance, and that estimate of variability can be obtained from the minimum design.

The error term "among areas within a zone" variance is inappropriate. To use this term would require an assumption that the study would use a random sample of areas within zones. In analysis of variance terms, areas would be considered random effects. In practice there are a limited number of areas which can be sampled and the spatial location of these areas or sub-populations are fixed (e.g., discrete riffle habitats within a zone). Because there is no element of randomization, the areas actually represent a fixed effect and the appropriate error term is "among fish" not "among areas".

Regardless of the sampling design, the difficulty with the EEM survey design is not in the statistical calculation but in the logical interpretation. Since there is no random assignment of treatments to zones we cannot be sure whether an observed difference is due to the "treatment", or to some inherent difference between the zones. One cannot confidently infer that the effluent is the cause of an observed difference. This logical problem is one shared by many epidemiology and sample survey studies, two populations are defined and random samples drawn from each; but because the units could not be assigned at random to the "treatments" it is impossible to be sure that the observed difference is actually due to the factor of interest. In cases such as these the inference of a cause-effect relationship requires ancillary information (e.g. demonstration that the effluent contains some substance which is known to affect the units) as well as information from appropriate follow-up studies.

Minimum Sample Size (20,20)

To obtain an estimate of the true standard deviation (SD) of a population, 20 observations is justifiable because there is little improvement in our confidence of the estimated SD with increased sampling effort. This is shown by Figure 5.2 which illustrates the changes in upper and lower 95% confidence limits of example SDs of 5 and 10, in relation to sample size (or *n*). The confidence limits are not symmetrical about the true standard deviation because the sampling distribution of variances is not symmetrical: it follows a Chi-Square distribution (Zar 1984). The equations for confidence limits about an estimated SD are:

Lower Limit: $$\sqrt{((v \bullet SD^2)/\chi^2_{\alpha/2,v})} \quad [1]$$

Upper Limit: $$\sqrt{((v \bullet SD^2)/\chi^2_{(1-\alpha/2),v})} \quad [2]$$

where,

v = number of degrees of freedom,
SD = estimated standard deviation,
χ^2 = Chi-Square statistic, and
α = probability that the true standard deviation (σ) is not included between the limits.

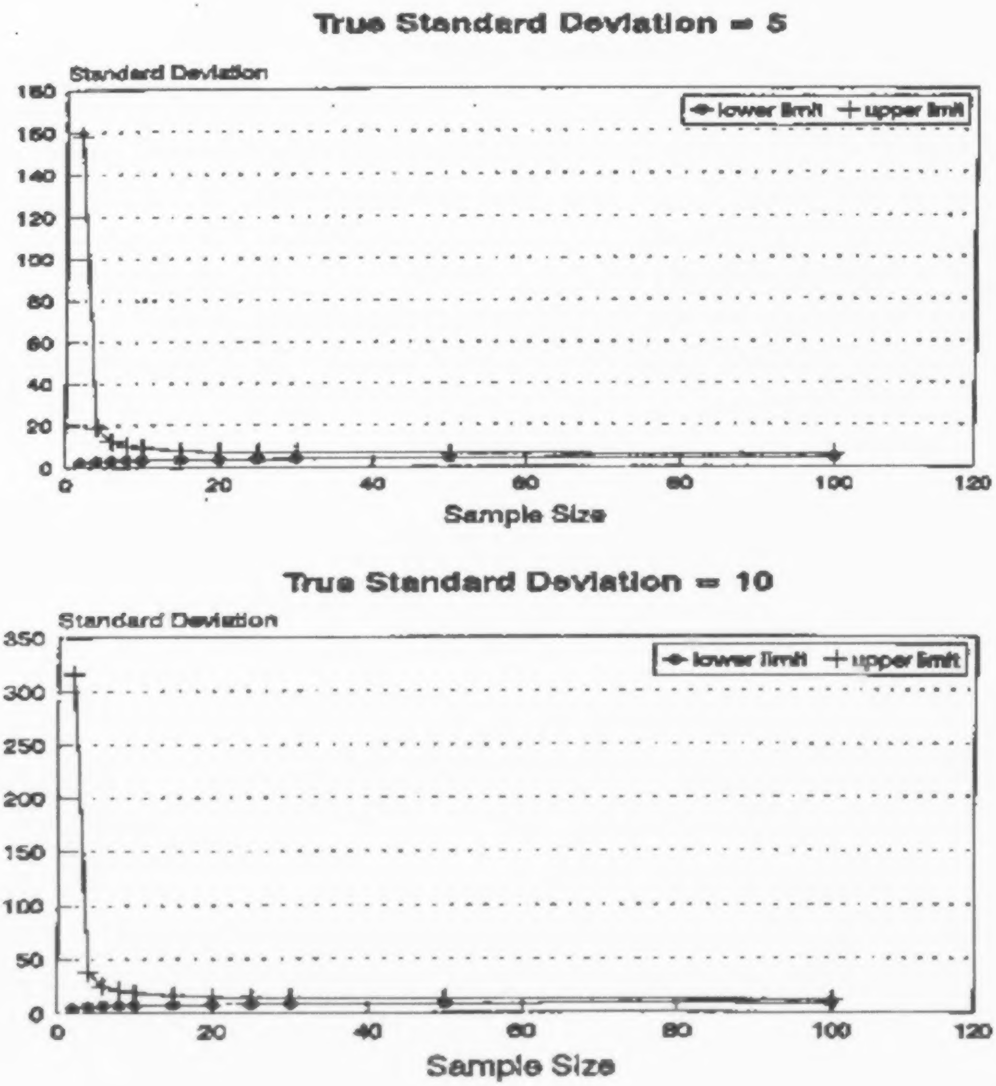
See Zar (1984; pp 114-115), Snedecor and Cochran (1980; pp 74-75) for more information on calculating confidence limits about SDs.

As Figure 5.2 indicates, there is little change in the 95% confidence limits with increasing sample size beyond a sample size $\approx$ 20. Note also that the there is no difference in the shape of the confidence interval curves for SDs of 5 and 10: the values change but the curves do not. Using data for white sucker, Munkittrick (1992) suggested that there was little improvement in variance estimates with sample sizes above 16 fish. This observation simply followed what would be expected from sampling theory and will apply to all EEM surveys regardless of the fish species used.

5.1.3.2 Timing of Collection

Fish should be collected during the period of residence in the effluent field. For species having spawning movements that take them away from or temporarily into the effluent field, a survey during the spawning season would be ineffective. Thus, for spring-spawning freshwater species, a fall survey would be appropriate. For fall spawners, a spring or summer survey is appropriate. This may not apply to fish in which ova mature rapidly, for example some late-spring-spawning minnows should be sampled in early spring, rather than in fall when ova may still be immature. Other factors to be considered in scheduling fish collections include water conditions (flow, turbidity, wave action), accessibility, and the cycle of gonadal development for the sentinel species. Some species undertake a very short period of gonadal growth and sampling schedules must optimize both residency and gonadal development concerns.

Figure 5.2 Changes in upper and lower 95% confidence limits of example standard deviations of 5 and 10, in relation to sample size.



5.1.3.3 Selection of Sampling Areas

Selection of sampling areas is one of the most critical components of the study design and should be considered carefully to maximize the amount of information gained from the survey. Definition and delineation of the exposure and reference areas and a description of the fisheries habitat are needed to locate the sampling stations.

The exposure area should be selected first to ensure that fish are obtainable there. The reference site can then be selected to provide fish of the same species, maturity, etc. as are available at the exposed site. Reference areas are ideally located in the same water body as the site discharge, but removed from any influence of the discharge (i.e. not exposed to site effluent). In general, reference sites should be:

- as similar as possible in all respects to the exposed area except for the effluent from the mill
- immediately adjacent to or upstream of the exposure zone unless that area is unsuitable (e.g. reservoirs, etc.)
- sampled at the same or appropriate time period
- described in as much detail as possible (physical, chemical and biological habitat of exposure and reference sites including measurements of temperature, depth, flow)

Selection of a reference site in Cycle One should not necessarily dictate its use in future cycles. Selection of reference sites should depend on:

- the mobility of the sentinel species (may eliminate adjacent or immediate upstream sites as reference sites)
- the timing of sampling (fish movements may render reference sites inappropriate during some seasons)
- sentinel species present in the exposed area; there is no purpose in collecting fish from a reference area if that species is not available in sufficient numbers in the exposed area.

Areas of concern in selecting reference sites are:

- presence of dams (can cause significant habitat changes which can alter the performance of some fish species)
- mobility of sentinel species
- confounding factors (e.g. other municipal or industrial outfalls); their presence should not automatically cause a rejection of the Fish Survey but the situation should be evaluated on a site-specific basis to determine the best approach
- natural variability

Multiple reference sites should be used rather than increasing sample sizes at one site, as a means of comparing natural variance among sites (which can then be compared to the variance between reference and exposed sites; see Section 5.1.7 for a discussion of sample variability).

Use of a general data set from regional reference sites is not acceptable because:

- of the difficulty of matching the receiving water characteristics of several mills within a region
- differences in the timing of sampling could have dramatic effects on interpretation
- it would generally require more reference locations than the traditional approach.

Combining benthos and fish sampling is not an acceptable criterion for selecting fish sampling sites because the characteristics of the fish species selected, the mobility of the fish, different habitat selection and the type of sampling gear required may make the optimal benthos sampling sites inappropriate for the Fish Survey. Although the sites may be the same in many circumstances this should not be a significant criteria for fish sample collection. However, when possible there are advantages to selecting similar sites so that the data can be used to help interpret responses.

Every attempt should be made to select reference and near-field areas that are as similar as possible in all characteristics except the presence of effluent. It is recognized that the two areas are unlikely to be identical, however, the assumption is that the differences in natural characteristics will be small relative to the presence of effluent. If this is not true, then this will become apparent in subsequent cycles.

Modifying/Confounding Factors

Variability in fish parameters (e.g. growth, weight, condition, etc.) is a function of a number of factors (modifying factors), not all of which will be related to the discharge of effluent. The FS design, by controlling for seasonality, habitat type and effluent exposure, attempts to minimize among-fish variance. Nevertheless, other natural or anthropogenic factors may modify fish parameters, increasing among-fish variability, or confounding the survey design.

If the range in values of modifying factors is similar among sampling areas, then effects of modifying factors can be considered negligible. However, when there are significant differences in modifying factors among sampling areas, the survey design becomes confounded. In this case, it is impossible to differentiate the effects of mill effluent from the effects of the modifying factor(s) on the response variable. For example, if the habitat type (e.g. pool) downstream of the mill was different from the habitat found upstream (e.g. riffle), then the effects of habitat type on fish parameters would confound any effects related to mill effluent; both mill effluent and habitat type may be good predictors of any differences in fish parameters observed upstream and downstream of the effluent discharge.

Examples of modifying/confounding factors to consider include:

- tributaries
- other point and non-point source discharges (e.g. industry, agricultural run-off, aquaculture practices, sewage treatment plants)
- natural environmental/habitat variables.

Tributaries

Tributaries provide dilution water to a main channel, lake, estuary or open ocean. This dilution water may or may not have similar chemical properties as the main channel under study depending on the similarities of the underlying bedrock. Tributaries also take time to mix with the main channels or receiving water body. As such, it is important to select sampling locations that avoid tributary mixing zones.

Other Point and Non-Point Source Discharge

Other point and non-point source discharges make it difficult to distinguish between effects due to mill effluent and other discharges, particularly if they are found within close proximity to the mill effluent discharge. In these cases, it will be impossible to avoid a confounded field survey as designed for this component of the EEM program. When other discharges are present immediately above the mill, the reference area is effectively limited to the area between the "other" discharge and the mill effluent discharge. In this way it may be possible to account for the influence of the other discharge, and to determine the magnitude of variance in fish parameters due to mill effluent.

Other discharges present in the study zone may be considered confounding if criteria established by the TMC and the FS EWG are met.

Natural Environmental/Habitat Variation

Using river systems as an example, the natural physical, chemical and biological properties change with distance downstream (River Continuum Concept; Vannote *et al.*

1980). This is a direct result of changes in river flow velocity, sedimentation, temperature, morphometry, etc. The physical changes result in changes in the biological components of the river. If study areas are extensive, it is possible that natural biological communities and characteristics will be different from one location to another. The influences of mill effluent, if any, and natural variation may be difficult to distinguish.

Naturally occurring modifying factors are also prominent in lakes, estuaries and marine systems. Examples of common, naturally occurring, and sometimes confounding, factors include:

- habitat type (riffle, run, pool)
- substrate type (organic content, particle size)
- water depth
- water flow rate/discharge
- tidal action/currents/wave exposure
- salinity
- dissolved oxygen/temperature
- emergent/submergent vegetation cover
- water chemistry (conductivity, hardness, pH, etc.,)
- biological properties.

Once present in the study design, confounding effects cannot be eliminated. Only by giving careful attention to potential modifying factors identified during the pre-design phase or the previous cycles of the FS can the influence of modifying factors be removed or controlled in subsequent cycles. Where it is not possible to eliminate confounding factors, increasing the number of sampling areas may allow the investigator to assess their influence on data interpretation (e.g. habitat of near-field area is different from the habitat of far-field area. It would be beneficial to sample separate reference areas corresponding to each of the two exposure habitats).

5.1.3.4 Fish Sampling Methods

Sampling methodology should be designed site-specifically, and capture gear and effort should be focused on methods shown to be successful. Standardized sampling is a priority. Therefore, in situations where sentinel species are the same as for a previous cycle, and the sampling techniques used previously were sufficient to capture the target number of each sentinel species, these same sampling techniques should be retained unless good reasons for change, such as unacceptable by-catch, are documented.

Gillnets

Gillnets may be appropriate in rivers and lakes. The correct mesh size should be selected to catch fish of the appropriate species and size, and to limit unwanted bycatch. Gillnets must be checked frequently to free unwanted fish. Fishing may be more successful at night, depending on the species sought.

Gillnets are not recommended where there is unacceptably high bycatch. There also can be major logistical problems with this technique because tidal or river currents can cause twisting of the nets. If gillnets are chosen as a technique, they can be sunk to fish at particular depths or floated to sample the surface layers only. Because the pelagic zone in marine receiving environments is typically stratified by salinity and temperature, the fish communities below and above the halocline are usually quite different. If the gillnet is several meters deep, the net will sample both below and above the halocline and the exact habitat where a particular fish species was caught may be difficult to determine.

Bycatch

It may be possible to obtain samples using the bycatch of commercial, research or other fisheries operations in either marine or freshwater situations. The investigator is responsible for ensuring and documenting that sampling procedures and conditions are met (QA/QC), and that fish are exposed. Capture techniques have to be standardized between sites.

Electrofishing

This method is useful in shallow rivers and lakes and may utilize a backpack or boat-mounted shocker. Unwanted bycatch can be released. Procedures that may enhance success include operating at dusk or night, passing over the same area at least three times, and using intermittent pulses of current (since a continuous field may actually chase fish away). It is important to ensure that the captured fish are representative of the population. Daylight sampling of shallow nearshore areas does not always product representative samples.

Trawling

Trawling may be appropriate for collecting bottom fish and some bivalves in coastal areas with a fairly flat seabed. On coastlines characterized by steep bathymetry (e.g. BC fjords), diffusers often discharge at depths greater than 20m and so introduce effluent into a steep underwater landscape. Trawling in these areas can be difficult because of the steep angle of the bottom. In these circumstances keeping an otter trawl open using the spreading forces of the otter boards may be difficult and beam trawls which use wood or metal frames to keep the net open may be preferable.

Underwater debris such as logs and rocky outcrops can also complicate trawling in the fjords and heavy mesh trawls resistant to tearing from snags should be used. The mesh size of the trawl is dependent on the species and size of fish targeted in the survey. Because changes in depth and substrate may result in changes in species composition of fish, it is important to ensure that the beginning and ending of trawls are precisely located using differential GPS. This should ensure that the same habitat is sampled in repetitive surveys. Other important variables such as time of start and stop of the trawl, depth, and substrate should also be noted.

Beach Seines, Fyke Nets, Trap Nets, Hoop Nets and Fish Traps

When effluent is discharged into rivers, shallow shorelines or intertidal habitat, seining may be considered. Permanent markers such as small cairns should be set up to mark the seining locations. Depending on the steepness of the shore, a small boat or persons wearing chest waders may be used to pull the net off the shore. Beach seines are routinely used for sampling small fish in rivers and estuaries and on beaches. On the Atlantic coast, fyke nets set in the intertidal zone are also an option. Baited traps are also used to catch mummichogs in Atlantic estuaries. Crustaceans such as prawns and shrimp can be caught with a variety of traps.

Hook and Line and Hand Collections

To minimize the impact of fishing gear on habitats, hand collection may be suitable for relatively sedentary species. This type of sampling can be very selective so that only the target species is sampled. For example intertidal blennies and cottids can be collected by searching under rocks and algae at low tide. Bivalves can also be collected by hand. Rockfish can be collected using SCUBA, a technique which allows sampling to 30m. Other selective methods, such as a string of baited hooks (longlining) for bottom fish such as sole, should also be considered. As with the other techniques precise positioning is vital to make sure the exact habitat can be revisited in the subsequent cycles.

Purse Seines and Midwater Trawls

Purse seining and midwater trawls may be used to sample species that live primarily in the pelagic zone. In order to effectively sample fish from depths where pulp mill effluent may be influencing fish living near bottom or in deeper waters, a fairly deep seine (e.g. 10 m deep) may have to be used. A large vessel such as a commercial salmon or herring seiner may have to be deployed. Midwater trawls have an advantage over seines in that the they allow more precise layers of the water column to be sampled and in addition the depth that the net is being towed can be recorded and/or monitored electronically and adjustments can be made as needed.

Remote Sensing

Video or still camera recording of fish abundance near outfalls can be monitored using cameras mounted on remotely operated vehicles. This technique may be particularly effective in rocky and steep areas where use of fishing gear may be difficult. Camera surveys may also be useful in reconnaissance surveys of bottom conditions before trawls or traps are deployed for fishing.

5.1.3.5 Sampling Effort

To aid in assessment of expected effort requirements at individual sites, the study plan submitted to Environment Canada must include all details on how fish sampling will be performed.

The following are performance-based criteria and guidance to determine a "reasonable level of fishing effort". Each site is unique. It is uncertain whether fishing success will be achieved at a site just because a certain level of sampling effort has been successful in the past at other sites or even at the same site at other times.

1. **The study plan** must include all details on how the adult fish sampling will be performed to aid in assessment of effort. Details to include in the study design (where applicable) are:
   - type and dimensions of gear (e.g. gillnet, trapnet, hoopnet, fish trap, trawl. In some cases more than one type of gear may be advisable.)
   - mesh type (e.g. nylon, cotton fibre or wire, knotted or knotless) and size
   - sampling time (i.e. time of day)
   - sampling duration (i.e., time interval between gear placement and retrieval)
   - frequency of checks:
     - ◆ location(s) of gear
     - ◆ how and why the sentinel species were selected
     - ◆ who was consulted on the locations and techniques chosen to collect the proposed sentinel species

- contingency plans regarding alternate gear and sentinel species
- scheduling of dates work will be performed so EEM contacts can be available for consultation.

Any preliminary fish survey results or observations made during pre-design activities should be provided where they have guided selection of sentinel species or procedures. The regional EEM contact will review these data and may request further information to clarify sampling procedures.

In some cases, the authorized study plan will specify a maximum level of fishing effort on a site-specific basis that has been jointly decided by the mill/consultant and the regional EEM contact.

2. **Proper operating procedures** must be used. Proper operating procedures include use of gear as outlined in the study plan. Gear must be checked at a frequency that ensures the recovery of sentinel species in useful condition and the release of non-target (especially protected and endangered) species. The use of non-lethal and/or selective techniques is a consideration here. A record of the identity and estimated numbers of non-target fish may be a useful addition to contingency plans. The mill and consultant must have a good understanding of the habitat, the characteristics of the sentinel species, and the gear being considered.

3. **Consultation with local experts** (e.g. provincial and federal fisheries personnel, individuals and associations involved in local sport and commercial fisheries, the public, and others with knowledge of local fisheries resources) must be demonstrated to ascertain that the selection of sentinel species, the location of nets, timing of collections, etc., are optimum.

4. **Personnel** charged with the fish collection and sampling procedures must have documented experience.

5. **Licence(s)** for collecting must be obtained from local fisheries agency(ies).

6. **Records** must be kept that document the operating procedures used (e.g. mesh size, sampling time, location, frequency of checks, etc.). These records may be required for review prior to a decision by Environment Canada.

Although not required, it is recommended that an estimate of catch per unit effort (CPUE) be provided for each sampling area (e.g. number fish caught per unit of time or area or net, etc.). The CPUE information is useful in documenting the effort expended in situations where collection of the minimum number of fish may be difficult.

Consultation with Regional EEM Contacts and Implementation in the Field

If all of the above criteria are met and a mill/consultant is having problems meeting the minimum data requirements of the adult fish survey, we recommend they consult with their regional EEM contact after (3) calender days or at consultation dates set between the mill/consultant and EEM contact to discuss possible solutions. If the EEM contact is unavailable then contact an alternate regional EEM representative (to be designated by regional EEM contact) or exercise options below using best professional judgment and discuss with the EEM contact as soon as available.

Possible outcomes and options of the consultation with the EEM contact are:

1. **Continue** - Advice on the following situations will depend on site-specific conditions. Set further consultation date(s) if required.
    a) Absence of target species at reference area:
        - continue with current gear and technique
        - continue at alternate reference area identified in contingency plan
        - continue at existing reference area with alternate target species, gear and/or technique identified in contingency plan.

    b) Absence of target species at near-field area:
        - continue with current gear and technique
        - continue at alternate near-field area identified in contingency plan
        - continue at existing near-field area with alternate target species, gear and/or technique identified in contingency plan.

    c) Absence of target species at reference and near-field areas:
        - continue with alternate areas, target species, gear and/or technique identified in contingency plan.

2. **Not to continue** at this time (influenced by dangerous conditions; sampling conditions (weather, cold) will not allow collection of fish; alternate gear not available; no further contingencies available (e.g. no further alternate sentinel species, further investigation needed)):
    - design new sampling plan in consultation with regional EEM contact
    - redeploy at a later date with original or alternate areas, target species, gear or technique but under more favorable conditions.
    - set date(s) for further consultation.

2. **Discontinue** - If the full complement of fish were not obtained, then the absence (or paucity) of fish will be considered a result which will be thoroughly explained in the study findings, taking all possible contributing factors into account. If the minimum number of fish is not caught, this could result in inflated variance estimates.

The decision to continue or not continue will be influenced first by safety considerations. In all scenarios refer to the contingency plan, where appropriate, and set date(s) for further discussion. Field technicians should speak directly with regional EEM contact. Sentinel species choices must apply to both reference and near-field areas. Pooling of data from different seasons is not valid.

Examples

Some examples of fishing methods that have been successful in collecting fish in a timely manner are provided below. These are provided as examples to guide consultants in the development and implementation of their study design, and to indicate when it may be advisable to consult with the Environment Canada regional EEM contact.

1. Nearshore Community Index Netting (warm freshwater fish communities): Data from 2 lakes indicated that 40 individuals of any of 6 species were collected in from 1 to 6 sets of 24 hours duration. The equipment included a 6 x 6 foot trapnet. Details of the process are in FAU Update Issue 94-1.

2. Spring Littoral Index Netting (cold freshwater fish community): Data from 7 lakes indicated that the fishing effort required to collect 40 lake trout varied from approximately 12 to 120 hours. The equipment included a 46 m gillnet gang with three 15.2 m panels. Details of the process are in FAU Update Issue 94-2, however, mesh size should be consistent and selected according to the target species.

3. Gill nets set in freshwater systems: Experience has shown that with 4 panels of 50 meters of net, set 24 hours a day for 5 days (equivalent to 24,000 meter-hours of effort) 20 fish of each sex can be collected. This is contingent on correct deployment of the panels and shifting of the panels to cover areas inhabited by the fish. Mesh size should be consistent and selected according to the target species.

4. Marine situations: Experience has shown that 48 hours of beam trawling, long-lining using a variety of hook sizes or other methods, including traps, (alone or in combination) should provide the 20 fish of each sex required.

5. Electrofishing (large rivers using boat-mounted apparatus): Consultation with users of this technique indicates that all fish can be obtained in one day. Procedures enhancing success include operating at dusk or night, passing over the same area at least three times, and using intermittent pulses of current (since a continuous field may actually chase fish away).

6. Bycatch: It may be possible to obtain samples using commercial, research or other fisheries in either marine or freshwater situations. The consultant is responsible for ensuring and documenting that sampling procedures and conditions are met (QA/QC).

Although the above techniques and gear may apply to a variety of species, these examples are not all inclusive since each site is unique and are provided as suggested effort only. Local expertise can provide further advice.

5.1.4 Alternatives to the Fish Survey

Given the variety of problems and situations encountered at each mill, there may be cases where the standard fish survey described above is not suitable. Alternative methods are under development and may be available to mills, where they are deemed appropriate by the Regional Authorization Officers, consultants and the mill, and they have passed the Screening Criteria developed by the Technical Management Committee.

5.1.5 Data Analysis

5.1.5.1 Variables to be Measured

For each adult fish collected, measurements are required of the following variables (see Table 5-1 for details):

1) length
2) total body weight
3) weight of liver
4) age
5) sex
6) weight of gonads
7) egg size/fecundity estimate for mature females

In addition, qualitative observations are to be made of the external condition of each fish, noting obvious abnormalities, prevalence of lesions, tumors and parasites, etc. A sample checklist for the physical examination of fishes is presented in Figure 5.3 and Table 5.10.

A list of appropriate aging structures for a variety of potential sentinel species has been provided in Table 5.9. Formal age validation is not a requirement of the FS; however, methods of aging must be consistent at each sampling area and among cycles and appropriate QA/QC procedures followed (e.g. independent confirmation). It is recommended that all aging structures be archived for future reference.

Figure 5.3 Features of composite fish.

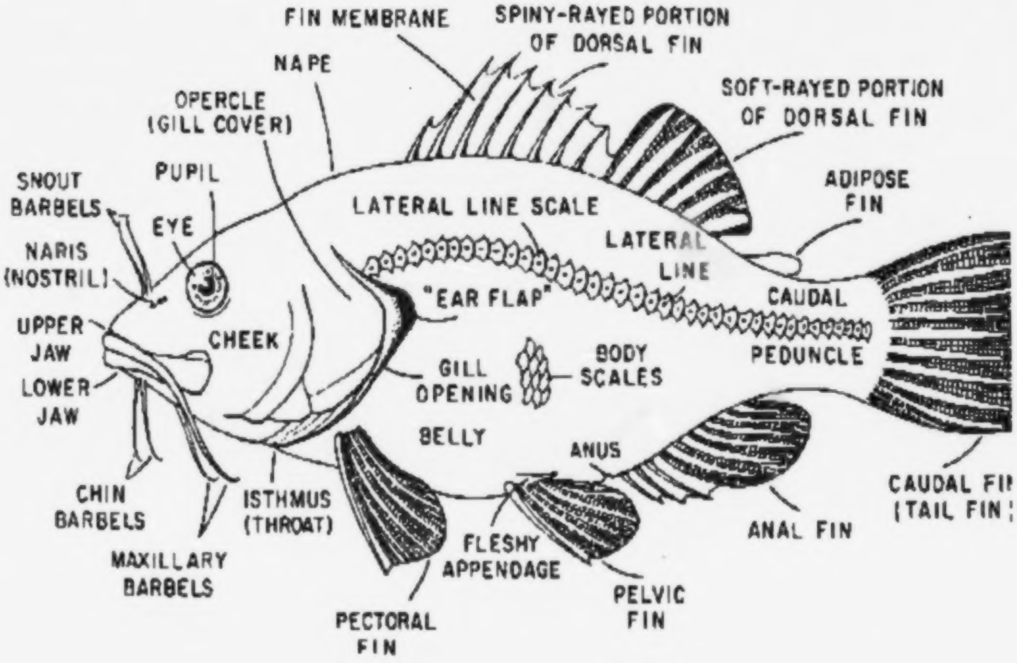

FIN MEMBRANE
SPINY-RAYED PORTION OF DORSAL FIN
NAPE
OPERCLE (GILL COVER)
SOFT-RAYED PORTION OF DORSAL FIN
SNOUT BARBELS
PUPIL
ADIPOSE FIN
LATERAL LINE SCALE
NARIS (NOSTRIL)
EYE
LATERAL LINE
"EAR FLAP"
CAUDAL PEDUNCLE
UPPER JAW
CHEEK
BODY SCALES
GILL OPENING
LOWER JAW
BELLY
ANUS
CHIN BARBELS
ISTHMUS (THROAT)
CAUDAL FI (TAIL FIN)
ANAL FIN
FLESHY APPENDAGE
MAXILLARY BARBELS
PELVIC FIN
PECTORAL FIN

Table 5.10 Sample Checklist for Physical Examination of Fishes[1]

Date:
Location:
Control no:
Sampling method:
Recorder:
Species:
Length (mm) Standard:
Total:
Fork:
Weight (g): Fresh:
Gonad:
Liver:
Sex:
Tissue samples:

BODYFORM

Normal
Emaciated
Truncate
Scoliosus
Lordosis

BODY SURFACE

Normal
Raised scales
Swollen
Lesions
Excess mucus
Reoriented scales
Growths
Parasites
Wounds (lamprey)

LIPS AND JAWS

Normal
Deformed
Growths

SNOUT

Normal
Pugnose (Pughead)

ISTHMUS

Normal
Enlarged
Hemorrhagic

EYES

Normal
Popeye
Cloudy cornea
Missing
Lens deformed
Lens parasites
Lens cataracts

FINS

Normal
Frayed - eroded
Parasites
Hemorrhagic
Gas bubbles

---

[1] adapted from Hunn, 1988

FINS-ERODED
- Dorsal
- Pectoral
- Pelvic
- Anal
- Caudal

BRANCHIAL CAVITY
- Normal
- Growths
- Parasites

LESIONS-location
- Fins
- Head
- Eyes
- Mouth
- Peduncle
- Ventral
- Dorsal
- Lateral

BARBELS
- Normal
- Deformed
- Missing

GILLS
- Normal
- Bright red
- Brown
- Gas Bubbles
- Parasites

OPERCLE
- Normal
- Incomplete

SEXUAL CONDITION
- Gravid

OTHER OBSERVATIONS

5.1.5.2 Data Handling

Upon completion of the field and laboratory measurements (including the appropriate QA/QC procedures), a data matrix with the location identifier (area), variables in columns, and observations in rows operates as the fundamental working unit. An electronic copy of the data will be required for submission to the Regional Authorization Officer. The initial step in data analysis is a check for transcription errors. Failure to identify these errors can invalidate further analyses. Assuming the data have been entered correctly, data should be summarized, screened for erroneous values and outliers, assessed for normality and transformed.

5.1.5.3 Data Summary and Screening

Visual screening techniques such as Box and Whisker Plots, Normal Probability Plots, and Stem-and-Leaf Diagrams can be used to identify extreme values (true outliers and/or data entry errors). Most statistical software packages provide data summary modules capable of generating appropriate summary statistics and graphics. It is also beneficial to summarize the data by calculating the mean and standard deviation (SD) of each variable separately for the reference and exposure areas. These summary statistics are usually needed for presentation and aberrantly high or low values can indicate errors. Extreme values or outliers should not be removed from the data set (unless they are obvious sampling, measurement or data entry errors) (Grubbs 1969; Green 1979) because this will result in the loss of statistical power to the FS. Instead, extreme values should be identified in the report and the influence of the extreme value(s) on the results should be determined by re-analyzing the data without the extreme value.

5.1.5.4 Assessing Normality and Equality of Variance

Data are typically tested for normality and homogeneity of variance before applying parametric procedures. It should be noted, however, that most of the univariate normal distribution-based statistical methods are very robust under moderate violations of the assumptions, with the exceptions of very small (<10 d.f. for error) and unequal sample sizes and/or one tailed tests (Harris 1975). For further guidance on issues related to assumptions and data transformation, it is recommended that the investigator consult Green (1979; Section 2.3.9, pp 44-63). Formal testing procedures for single variates have been described by Klemm *et al.* (1990) and Gibbons *et al.* (1993). Probit plots can be used to assess normality. If the plot is not visibly non-linear it is unlikely that the departure from normality is serious enough to cause problems.

Analysis of Residuals

The preferred method of examining the residuals is to use graphical methods rather than relying on formal tests to assess normality and equality of variance. In fact, Day and Quinn (1989) have recommended against using formal tests. A good discussion of this topic can be found in Miller (1986). Draper and Smith (1981) review various methods of examining residuals, particularly residuals from regressions. Most statistical software packages also provide modules for examination of residuals. These methods are usually graphical, although there are diagnostic statistics available as well. The primary advantage of these methods, compared to formal tests, is that they can identify the cause of violations of normality or equality of variances.

5.1.5.5 Data Transformation

Even when the statistical procedures are robust to violations of normality and/or equality of variances, it is often possible to reduce the violations through appropriate transformations of the original data (Green 1979). Taylors Power Law (Taylor 1961), which examines the relationship between treatment means and variances, can be used to determine the specific transformations to normalize data or homogenize variances (Green 1979). However, most biological measures are considered to operate on a log or exponential scale (Peters 1983) and, in practice, a logarithmic transformation is preferred over other transformations (e.g. square root, inverse, fourth root, etc.) because it is more comprehensible and biologically meaningful.

5.1.6 Fish Parameters

5.1.6.1 Principal Parameters and Estimation of Variability

The variables listed in Section 5.1.5.1 will be used to derive the following parameters:

- growth: size (both as length and weight)-at-age
- reproductive investment: gonad weight (as gonad weight @ body weight; gonad weight @ length), fecundity (as fecundity @ body weight; fecundity @ length), age at maturity
- age structure: mean adult age
- energy stores: condition (body weight @ length), liver weight (as liver weight @ body weight and liver weight @ length)

Sex differences in growth rate, body weight, gonadal size and liver size are common due to differences in overall energetic requirements between male and female fish. Therefore, for all parameters, sexes must be treated separately when estimating variability. Immature fish should not be mixed with mature fish for analyses.

Most of the parameters, with the exception of mean age and age at maturity, are estimated as regressions of one variable on another using Analysis of Covariance (ANCOVA) (Table 5.11). ANCOVA combines the features of analysis of variance and regression, and can be used to compare regressions among multiple classifications (i.e. sampling areas) (Snedecor and Cochran 1980). The mean squared error (MSE; model variance, s2 [= SD2 of the residuals from the regression]) of the ANCOVA model should be used to provide an estimate of variance of the fish parameters, when reasonably large samples are used.

Assumptions of ANCOVA are that:

- the slopes of regression lines among areas are equal
- the independent variable (covariate) is fixed and measured without error
- the relationship has the form specified (linear regression)
- the residuals are normally and independently distributed with zero mean and a common variance.

The assumption that the independent variable is fixed is frequently violated and the consequences of this violation are discussed by Draper and Smith (1981). In the case of the FS, this is likely to prove problematic only in situations where the range of the independent variable is very small. The assumption of a linear relationship can be tested for samples with multiple observations at different values of the independent variable. This may be possible for discrete variables such as age but not for continuous independent variables such as body weight.

Some workers suggest that the range of the independent variable should be approximately the same for each ‘treatment’ (area). This will be difficult to assure in practice, but the violation of this should be considered when interpreting results from such cases. Some also consider a further requirement to be that the range of the independent variable be ‘compact’. This serves to avoid difficulties created by outlying values of the independent variable if there is some non-linearity in the relationship.

A brief description of the application of ANCOVA for determining size-at age follows to provide a better understanding of the procedure.

Size-at-age

Rates of growth are commonly described by the relationship of size (as weight or length) to age. Over the entire life-span of a fish, this relationship is curvilinear, with the rate of increase declining as fish approach the limit of their life span (Ricker 1975). Because the standard FS samples only adult fish, classical growth rates cannot be calculated. Nevertheless, for the purposes of the FS, fish growth can be inferred from size-at-age estimates determined for each area using ANCOVA. This calculation assumes that the relationship between size and age for adult fish is approximately log-linear (log size vs log age) (Bartlett *et al.* 1984).

Size-at-age may be estimated by calculating the regression relationship between body size (weight or length) and age for each sampling area (reference and near-field). It is recommended that both length and weight be used to calculate size-at-age in order to determine which provides the best fit and tightest regression. Analysis of covariance (ANCOVA) can then be used to compare the separate regressions between body weight (dependent variable; Y) and age (covariate; X).

Table 5.11 Definition of fish parameters used in the Fish Survey (from Gibbons *et al.* 1992).

Parameter	Dependent Variable (Y)	Covariate (X)
Age	Age (log)	none
Size-at-Age	Body weight and length (log)*	Age (log)
Condition	Body weight (log)	Length (log)
Gonad weight	Gonad weight (log)	Body weight and length (log)*
Fecundity	Number of eggs/female (log)	Body weight and length (log)*, and age
Liver weight	Liver weight (log)	Body weight and length (log)*

* Both weight and length should be used as an estimate of size when calculating growth rates, and used as covariates when analyzing gonad weight, fecundity and liver weight (see text).

As previously mentioned, an assumption of ANCOVA is that the slopes of the regression lines among areas are equal. It is first necessary, therefore, to test for differences in slopes between areas (described by the interaction term of age*area from ANCOVA). In cases where the slopes are not significantly different, this indicates that the regression lines are parallel to each other i.e. that weight gain over age is similar for both areas. The next step is to proceed with the ANCOVA model and test for differences in intercepts (elevation) to investigate whether fish are proportionately heavier at any age in one area than in another. Although uncommon, it is possible that the slopes of regressions may differ, i.e. fish from one area are gaining more weight with age (steeper slope) than fish from another area. In this situation, it is a good idea to plot the separate regression lines to obtain a better understanding of the size-at-age relationship at each area, over the entire data range (e.g. where do the lines intersect?).

The objective of a first cycle EEM Fish Survey was to derive an estimate of variance for parameters which is best obtained from the MSE of the ANCOVA model. In cases where the assumption of equality of slopes was not met, and the range of ages are similar among areas, then it is not necessary to compute the MSE as the minimum sample size has been demonstrated to be sufficient to detect differences between areas.

Condition, fecundity, gonad weight and liver weight

A similar procedure as outlined for size-at-age can be used to estimate the variability in fish condition between areas. In this case, condition is defined by the relationship between body weight and body length (Table 5.10) and essentially describes how "fat" fish are at each area. For fecundity, gonad weight, egg weight and liver weight, the same procedure is appropriate and body weight and length, (see above discussion on calculating size-at-age) are used as the covariate to adjust for any differences in size in order to place the variables on a relative scale (Table 5.10).

Mean age

Calculation of mean age is meant as a gross reflection of the age distribution of adult fish collected from each area. Variability in mean age of fish can be estimated using Analysis of Variance (ANOVA). The MSE from the model is the best estimate of variability.

Age-at-maturity

Age-at-maturity is a commonly used parameter in fisheries science, however, few methods of calculation incorporate a measure of statistical confidence or variability. Therefore, it is recommended that age-at-maturity be estimated by traditional probit analysis as commonly used for determining LC50 in toxicity tests. By determining the proportion (%) of mature individuals in each adult age class, and converting these data to probits (or plot the data on probit paper), a straight line relationship is generated (probit vs. log age) that allows one to estimate the age where 50% of the fish sampled are sexually mature. An estimate of variability in age-at-maturity among individual fish can be obtained from the slope of the line. The slope estimates 1/standard deviation (SD).

Therefore, the standard deviation is estimated by 1/slope. In future cycles, confidence limits can be calculated as an estimate of precision and statistical comparison of area values. Most statistical software packages can convert percentages to probits and there are several small independent packages designed to conduct LC50/probit analysis and generate the confidence limits. For more detailed information on conducting probit analysis, refer to Hubert (1980). For a discussion on factors to be considered when using probit analysis, as well as other techniques, for estimating age-at-maturity, refer to Trippel and Harvey (1991).

Owing to the preliminary nature of the FS in Cycle One, and in the absence of existing data, providing an estimate of variance for age-at-maturity will likely be opportunistic for some populations, i.e. all fish collected from an area may be mature or immature. This is especially true for collection methods that are very size-selective (e.g. gill net). In these situations, it is not possible to estimate age-at-maturity (or associated variability). Sampling in subsequent cycles can be designed to bracket the age/size range of fish at which attainment of sexual maturity is expected.

5.1.6.2 Use of Indices

A variety of indices have been used in fisheries biology to describe the condition of fish (Bolger and Connolly 1989). Many of these indices were derived by calculating the ratio of one variable to another. Examples of a few common indices are:

- condition factor (k)=100(body weight/length^3)
- gonadosomatic index (GSI)=100(gonad weight/body weight)
- liversomatic index (LSI)=100(liver weight/body weight).

In general, investigators have become cautious about using derived variables and ratios because they assume some relationship between variables which may not exist, and because they may have unusual and undesirable statistical properties (Green 1979; Cone 1989; Jackson *et al.* 1990). Although these indices may be used for presentation purposes, it is preferable statistically to estimate (and analyze) the parameters based on regressions of original variables (i.e. ANCOVA) rather than ratios (Gibbons *et al.* 1993).

5.1.7 Variability and Sample Size Considerations

Using power equations, the estimated variability in a specific parameter can be used to calculate the appropriate sample size (number of fish) required to detect a given effect size in that parameter. This information will be useful when developing the study design for future cycles.

For spatial comparisons with equal replication, the power equation from Green (1989) has the form:

$$n = 2\,(t_{\alpha(2)} + t_{\beta})^2\,(SD/\delta)^2$$

where:

n is the number of samples
SD is the sample standard deviation (or $\sqrt{}$MSE from ANOVA or ANCOVA model)
δ is the specified effect size (between reference and near-field area)
$t_{\alpha(2)}$ (two tailed test) and t_{β} (one tailed test) are critical values of the Student's t-test statistic (given Type I error probability α and Type II error probability β - taken from Student's tables with 2(n-1) degrees of freedom).

The above equation calculates exact sample sizes using t-values and solving iteratively. A simpler equation can be used which replaces t-values with Z-values (standard normal deviates) (Alldredge 1987; Snedecor and Cochran 1980):

$$n \geq 2\,(Z_{\alpha(2)} + Z_{\beta})^2\,(SD/\delta)^2 + 0.25Z_{\alpha(2)}^{\;2}$$

This formula is an approximation of the t-value equation and does not require an iterative solution. The term $0.25Z_{\alpha}^{2}$ is a correction factor which accounts for using Z-values rather than t-values. The remaining terms are the same as described for the t-value power equation.

A Type I error is the probability (α) of reaching a false positive conclusion of effect or difference when none exists, and is the risk to the mill. A Type II error is the probability (β) of reaching a false negative conclusion of no impact or difference when there is a true effect of magnitude δ, and represents a risk to the environment (or fish population). The power (1-β) of the comparison, for any δ, is the probability of detecting a true difference of some specified magnitude.

Given the target effect size (δ) of 20-30%, and the specified values for α and β, n can be solved to determine the sampling effort needed to detect the given effect size. Without information on costs of Type I or Type II errors, it is difficult to specify acceptable values of α and β. The approach used here sets $(1-\beta) = 0.80$ to maintain fairly high power (i.e. probability of detecting a difference will be 80%). By convention α is set at 0.05 and results in a conservative test that requires a higher number of samples than $\alpha = 0.10$, but provides less risk of a Type I error to the mill.

For more detailed information regarding power analysis, consult Skalski and McKenzie (1982), Bernstein and Zalinski (1983), Alldredge (1987), Green (1989) and Peterman (1990). Goldstein (1989) provides a review of computer programs for power analyses. The Fish Survey Expert Working Group Final Report also provides information regarding statistical sampling design and power analysis.

5.1.8 Reporting

The following section outlines a list of possible reporting considerations related to the Fish Survey.

Fish Survey QA/QC

An outline of recommended QA/QC requirements for specific components of the Fish Survey are presented below. The following QA/QC considerations are appropriate for the Fish Survey, and may also be applied to the next cycle of follow-up studies. Most of the following information has been taken from Environment Canada & Department of Fisheries and Oceans (1993) and Gibbons *et al*. (1993). This information focuses on study design, field sampling, sample processing and laboratory analyses, data analyses, and reporting.

Study Design

Before the survey is implemented, the study design should be reviewed and approved by the researcher, regulator and a knowledgable expert (e.g., biometrician). The review should evaluate:

- how well the design meets the study objectives
- the selection of appropriate sentinel species
- the statistical elements of the design (including statistical power if historical fisheries data is available)
- the technical elements of the design (e.g., sampling areas and location, collection frequency and timing, collection methods)
- the overall feasibility of the proposed design (technical, logistical, statistical, economic, etc.,) considering appropriate modifications or possible alternative designs (if needed) that would reflect the study objectives.

Field Sampling

This is the first stage of data collection. QA/QC procedures for the Fish Survey should be outlined during the development of the study plan and should be followed precisely in order to maintain high quality data. While a QA/QC plan for field sampling can have many components, some of the main procedures are as follows:

- initiate and maintain communication with local government agencies (e.g., fishing licence, dates of fish collection, location of collection, etc.)
- all personnel involved in field sampling should have appropriate education and/or training
- all safety measures must be identified, understood and adhered to
- fish collection methods and equipment must be appropriate for the specific waterbody and sentinel species.
- habitat descriptions including possible modifying factors (water depth and current, dissolved oxygen, temperature, substrate classification, evidence of pollution [discolouration, odour, residues], salinity, conductivity, etc.)
- date and time of collection
- collection methods need to be consistent throughout study
- location of sampling areas and fish collection sites documented; photograph collection location
- record of the number of sentinel species and incidental species caught per collection location
- estimate of catch per unit effort
- samples from fish (e.g., ovaries, age structures, stomach contents) must be placed in appropriate containers
- suitable preservatives/fixatives (e.g., ovaries - frozen or formalin, stomach contents - formalin) must be used
- all samples must have appropriate labelling
- all measurements must be taken using appropriate equipment of acceptable accuracy and precision (this should be documented)
- instruments should be calibrated and maintained in good working order (records and methods should be available)
- detailed field notes should be maintained in a bound notebook
- chain-of-custody forms and appropriate shipping and storage procedures should be used

Sample Processing/Laboratory Analysis

Although much of the survey information is collected while in the field, variables such as fecundity, egg size and age are usually determined later in the laboratory. With each measurement, the primary concern of the laboratory QA/QC program is to ensure consistency (precision) and accuracy of the data. The following issues should be considered as part of the measurement procedures:

- all personnel involved in sample processing and analyses should have appropriate education and/or training
- measurements should be conducted using recognized protocols and methods (these should be documented) all instruments used should be properly calibrated and maintained (records, methods available)
- record of fish measurements recorded for each fish (sentinel species)
- record of external lesions, tumours and parasites, etc.
- record of evidence of parasitism
- fecundity data including methods and subsampling precision (if applicable)
- ageing data including methods and independent confirmation of estimates
- records should be kept describing the sample, measurement, and responsible personnel. If possible, a minimum number of individuals should conduct a particular measurement to maintain consistency and reduce measurement error (especially for age determination)
- if subsampling is necessary (e.g., fecundity, egg size, stomach content analysis), information describing the efficiency and accuracy of the subsampling technique should be documented. This information should also be used to calculate appropriate correction or scaling factors (if needed) to minimize the affects of possible confounding differences in methods and efficiency. This is especially true for situations where different levels of effort are used to determine the value of a variable for different fish (e.g., sorting stomach contents - large, full samples may need to be subsampled before sorting, whereas it is feasible to sort 100% of smaller samples).
- all data should be verified. For example, measurements such as fecundity and egg size should be replicated to ensure precision and accuracy. Estimates of age should be verified by a recognized expert. Identification of stomach contents (particularly benthos) should also be verified by a knowledgeable expert. Sorting efficiency of stomach contents should be checked by re-sorting a representative sample (e.g., 10-20 %) to determine percentage of animals missed.
- literature, taxonomic keys used for benthos, and fish prey identification (stomach content analysis) should be documented
- archive samples and voucher specimens
- maintain detailed sample processing and laboratory notes in a bound notebook

Data Analysis

QA/QC concerns regarding data analysis include: data verification and validity, repeatability and robustness of statistical analyses, and rigour and defensibility of analyses. For the most part, validation and verification of data depends on the success of QA/QC procedures during field sampling, sample processing, and laboratory analyses (see above). However, there are other considerations regarding the data verification and analyses:

- conduct screening techniques to identify possible transcription errors, outliers and other "suspicious" data points,
- tabular summaries of the mean and standard deviation of fish measurements (see Table 5.12 for example)
- results of assessing assumptions of normality and homogeneity of variance
- record of transformation used
- parameter estimates of variability (ANOVA MSE, ANCOVA MSE, SD for age-to-maturity)
- calculated sample size requirements for each parameter
- summary of adherence to data quality objectives, standard operating procedures and identification of any QA/QC problems incorporate considerations related to laboratory and field QA/QC,
- to allow reproduction of analyses and results, provide all raw data in an Appendix and archive computer data files for an approved period of time after the analyses are published in a report,
- document in detail a clear description of the methods used for analyses,
- verify that statistical software packages used produce the same output and results as other packages,
- evaluate the robustness of the analyses, i.e. the results and conclusions should be similar
- regardless of whether outliers are included or excluded, transformations are used or not, etc. The objective is to ensure that results are not a function of some manipulation or assumption prior or during analyses,
- maintain detailed notes regarding the analyses of the survey data

Reporting

The report must document all information pertinent to the study and to the evaluation of the results. Reporting QA/QC should review:

- editorial, grammatical and spelling errors, data entry errors
- consistency of format and the completeness of each section
- data handling and reporting (entry checks, missing values, methods, QC data)

- all methods (collection, processing, laboratory, analyses)
- that pertinent information has been reported in detail (including field notes, map accurate site locations)
- changes in protocol, study design, or other components of the study
- accuracy of analysis, feasibility of recommendations
- all QA/QC documentation included in the report (as an appendix)
- all raw data in appendices
- the completed study report

Table 5.12 Suggested reporting format for the parameters (A) and the resulting regressions (B) required for the fish survey analysis. The percentage difference should be reported as exposed relative to reference site. Statistical significance should be given as p value.

A. Parameter Summaries

Species	**Sex**	**Parameter**	**Reference**		**Exposed**		**% Diff**	**Stat. Sign**
			Ref (n)	Reference Mean and Std. Dev.	Exp (n)	Exposed Mean and Std. Dev.		

B. Regression Analyses

Species	**Sex**	**Regression**	**Ref (n)**	**Reference Adj. Mean**	**Exp (n)**	**Exposed Adj. Mean**	**% Diff.**	**Stat. Sign.**	**Sign. Interax.**

5.2 Tissue Analyses: Chlorinated Dioxins and Furans Congeners

Mills which use or have used chlorine bleaching may be required to conduct analysis of tissue levels of chlorinated dioxin and furan congeners on edible portions of fish if dioxins and furans are an issue for the receiving environment (see Figure 5.4).

Analyses must be conducted unless it can be shown by the mill that:

1) Levels of dioxins in the previous cycle were below the consumption guidelines or there is no consumption guideline-related advisory/closure based on dioxin/furan contamination for any species or tissue in the receiving environment of the mill

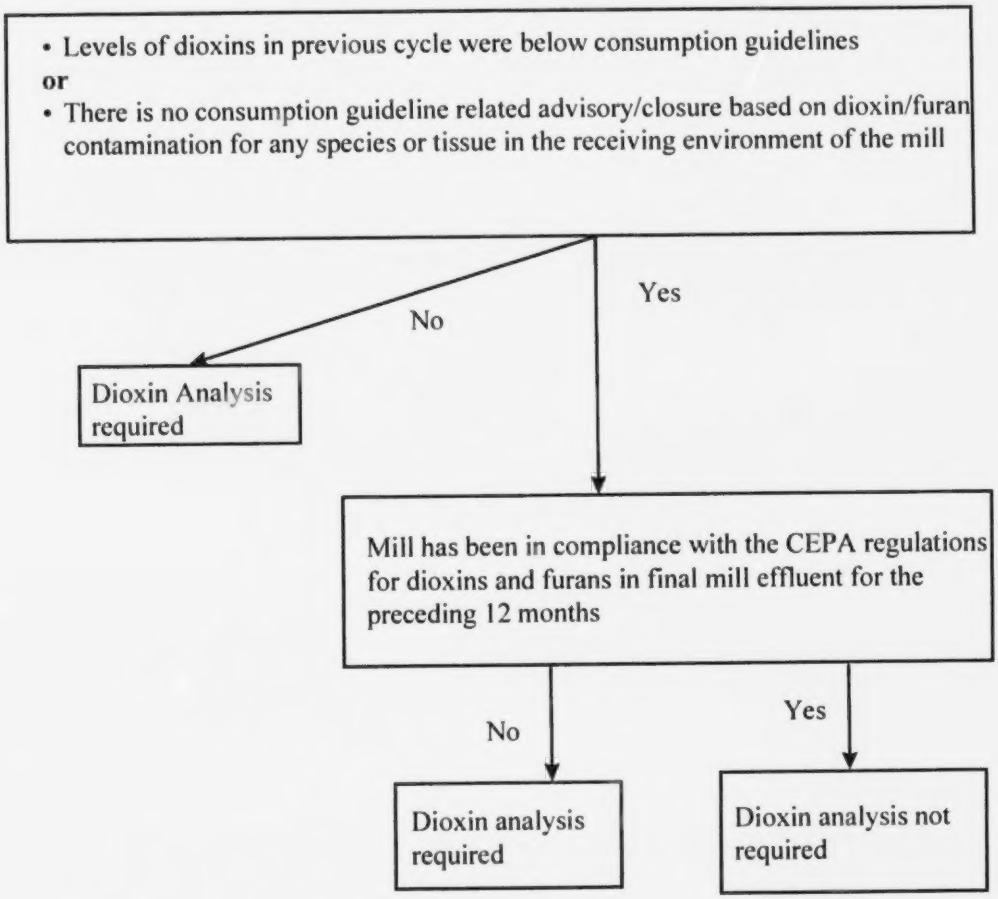
and

2) The mill has been in compliance with the CEPA regulations for dioxins and furans in final mill effluent for the preceding 12 months.

The species selected for analyses and that portion of the fish constituting the edible portion will be decided on a site-specific basis. At sites where the levels of dioxins reported are approaching the advisory guideline and there are no other supporting data or programs, the Regional Authorization Officers may require that dioxin analysis be included in the next cycle of EEM. Where mills are already engaged in monitoring programs to assess levels of dioxins and furans in fish tissue to conform with other federal or provincial requirements, then the existing program shall take precedence, providing it conforms with the performance requirements described in Section 5.2.2. Data collected for purposes other than the EEM may be substituted if they:

- were collected after the completion of the previous cycle, and
- meet the minimum QA/QC requirements as described in Section 5.2.2, and
- are equivalent to the sampling requirements (i.e. composite of 10 fish), and are reported with supporting data and information in the EEM document.

Figure 5.4 Dioxins and furans decision tree


• Levels of dioxins in previous cycle were below consumption guidelines
or
• There is no consumption guideline related advisory/closure based on dioxin/furan contamination for any species or tissue in the receiving environment of the mill
No
Yes
Dioxin Analysis required
Mill has been in compliance with the CEPA regulations for dioxins and furans in final mill effluent for the preceding 12 months
No
Yes
Dioxin analysis required
Dioxin analysis not required

5.2.1 Tissue Sample Collection and Preparation

Analyses will be conducted on a composite sample of 10 individuals of a single species and sex (male or female) from the near-field area and reference area. If possible, the composite should be of one sex, but if this is not possible, then the sex of each fish making up the composite must be reported. Fish collected should be adults. The species selected for analyses and that portion of the finfish constituting the edible portion will be decided on a site-specific basis. Emphasis should be made on using an important commercial or sport fish species. If available, lobster or crab must be sampled in marine areas. For lobster and crab, the tissue to be collected is hepatopancreas. If a commercial or sport fish/shellfish is not available, a species that has been historically monitored may be substituted. If an alternative species is used, a justification must be included in the report. The length, weight, sex and age (if possible) of the fish making up the composite samples should be similar between reference and exposed areas. These data should be reported.

Tissues collected for analysis should be handled to avoid organic contamination from sources such as survey boat fuel. For finfish, edible portions (fillet) should be removed, skinned and deboned in the laboratory, with a hexane-rinsed, stainless steel knife and frozen in hexane-rinsed aluminum foil. Each pooled sample should consist of a minimum of 50 g fresh weight of tissue to provide sufficient material. Each sample should be clearly labelled and forwarded to the analytical laboratory for determination of specified data.

For molluscs, whole soft body parts should be collected, and it may be necessary to composite tissues from more than ten specimens to create an adequate sample weight. For lobster or crab, edible tissue (muscle and hepatopancreas) should be collected Samples should be handled as outlined above for fish tissues.

A composite sample of at least 50 g should be provided to the laboratory. Composite samples must be made up using an equal wet weight of tissue from each fish. Samples are to be homogenized and a subsample taken for dioxin and lipid analysis.

5.2.2 QA/QC Criteria for Clorinated Dioxins and Furans Analyses

Performance criteria for analytical precision and accuracy for chlorinated dioxins and furans analyses are given in Table 5.13. Criteria for quality assurance and quality control for the conduct of clorinated dioxins and furans analyses are summarized as follows:

1. All dioxin/furan analyses must be performed by high resolution gas chromatography-high resolution mass spectrometry (HRGC/HRMS).
2. All data reporting must be done as per Environment Canada Report EPS 1/RM/23,

October 1992, Figure 3. An equivalent format is acceptable if all information is provided.

3. Instrument detection limits must be reported for at least the 2,3,7,8-TCDD congener.

4. Method detection limits (MDLs) must be reported for all the 2,3,7,8 PCDDs/PCDFs. Minimum requirement for MDLs are those postulated in the Environment Canada (1992b) (Report EPS 1/RM/23, October 1992), Table 1.

5. Composition of every sample batch that the submitted samples were processed in must be provided. Batches of 12 samples are preferable, and must contain a procedural blank, a CRM or an in-house spiked sample, one replicate, and 9 real samples.

6. Data for the glassware blank associated with each of the samples analyzed must be provided. It must also meet the Environment Canada (1992a) (EPS 1/RM/19) criteria, Section 8a on page 37.

7. The procedural blank must be submitted with each sample batch and must meet the Environment Canada (1992a) (EPS 1/RM/19) criteria, Section 8e on page 37. Contamination levels of individual 2,3,7,8-substituted tetra, penta, hexa, hepta, and octa CDD/CDF must not exceed 1.5, 3, 3, 5, 10 pg/sample respectively.

8. Each batch of 12 samples (as specified in #5) must have one sample analyzed twice (a replicate). Acceptable precision for replicate analysis must be better than ±15% for sample concentrations of 5 pg/g and higher and ±30% for sample concentrations lower than 5 pg/g.

9. The acceptable percent recoveries for the surrogate method internal standards must be those postulated in the Environment Canada (1992b) (Report EPS 1/RM/23, October 1992), Table 2. For OCDD surrogate recoveries should be between 35-120%.

10. Each batch of 12 samples (as specified in #5) should contain a certified reference material (CRM, or an in-house spiked sample) and an analytical report must be provided that outlines the measured concentrations and their % deviation/recovery from the expected concentrations. For congeners with concentrations higher than 1 pg/g the percent deviation must be less than ±20% from the expected concentration. For in-house spiked samples this will reflect the spiked concentrations and for CRMs this will reflect deviation from the mean of the consensus value.

Table 5.13 Performance criteria for analysis of chlorinated dioxins and furans

Variable	Matrix Detection Limit	Precision	Accuracy
2,3,7,8-substituted PCDD[2]			
2,3,7,8-TCDD	1 pg/g	± 20 %	40 - 120 %
1,2,3,7,8-P_5CDD	5 pg/g		
1,2,3,4,7,8-H_6CDD	5 pg/g		
1,2,3,4,6,7,8-H_7CDD	5 pg/g		
0_8CDD	20 pg/g		
Total PCDD[3]			
TCDD	1 pg/g	± 20 %	40 - 120 %
P_5CDD	5 pg/g		
H_6CDD	5 pg/g		
H_7CDD	5 pg/g		
0_8CDD	20 pg/g		
2,3,7,8-substituted PCDF[4]			
2,3,7,8-TCDF	1 pg/g	± 20 %	40 - 120 %
1,2,3,7,8-P_5CDF	5 pg/g		
2,3,4,7,8-P_5CDF	5 pg/g		
2,3,4,6,7,8-H_6CDF	5 pg/g		
1,2,3,4,6,7,8-H_7CDF	5 pg/g		
0_8CDF	20 pg/g		
Total PCDF[5]			
TCDF	1 pg/g	± 20 %	40 - 120 %
P_5CDF	5 pg/g		
H_6CDF	5 pg/g		
H_7CDF	5 pg/g		
0_8CDF	20 pg/g		

---

[2] 2,3,7,8-substituted PCDD means any polychlorinated dibenzo-para-dioxin that has the molecular formula $C_{12}H_{8-n}Cl_nO_2$ in which "n" is not less than 4 and not more than 8 and chlorine atoms are located at the 2,3,7,8 positions on the molecule:

[3] Total PCDD means the sum of the concentrations of all polychlorinated dibenzo-para-dioxin congeneres that have the molecular formula $C_{12}H_{8-n}Cl_nO_2$ in which "n" is not less than 4 and not more than 8;

[4] 2,3,7,8-sbstituted PCDF means any polychlorinated dibenzofuran that has the molecular formula $C_{12}H_{8-n}Cl_nO$, in which "n" is not less than 4 and not more than 8 and chlorine atoms are located at the 2,3,7,8 positions on the molecule;

[5] Total PCDF means the sum of the concentrations of all polychlorinated dibenzofuran congeners that have the molecular formula $C_{12}H_{8-n}Cl_nO$, in which "n" is not less than 4 and not more than 8.

5.2.3 Lipid Analysis

Tissue residues of chlorinated organics have been correlated with lipid concentrations in fish and invertebrates. Lipid concentration consequently has been used to normalize tissue residues among species or within species between seasons, as well as being a key variable in modelling bioaccumulation. Lipid extraction methods by Randall *et al.* (1991) and the chloroform-methanol extraction method developed by Folch *et al.* (1957) and modified by Bligh and Dyer (1959) are recommended.

Percent lipid and percent moisture determinations must be provided for every sample submitted for chlorinated dioxins and furans analysis. Also percent lipid values should be reported for the replicates analyzed in the same batch with the submitted sample. The percent lipid precision for the replicate samples must be ±30% for tissues containing more than 2% and ±60% for tissues with less than 2% lipid. The method for the lipid determinations must be reported and the solvents used must be clearly specified.

5.3 Tainting Evaluation

5.3.1 Introduction

This section addresses, in detail, all aspects of a program to assess impairment of the usability of fisheries resources within the context of a Pulp and Paper EEM program. It will assist pulp and paper mills and regulatory agencies in the design and undertaking of taint assessment studies that will provide scientifically defensible data to assess the usability of fisheries resources in situations where impairment of the resources is suspected. Special attention has been made to the use of the terms "should" and "must". The final EEM reports must document that the technical requirements that are preceded by a "must" have been satisfied. Requirements preceded by "should" are not absolute requirements but meeting these requirements will support the overall technical quality of the data.

The definition of tainted fish used in this document is based on the ISO 5492: 1992 Sensory Analysis Vocabulary and is defined as "fish that contains an abnormal odour or flavour". The context of the definition is that the odour or flavour is introduced into the fish from outside sources and does not include any natural by-product from the deterioration in quality due to ageing of fish in storage, decomposition of protein, fat, or other components, or due to microbial contamination normally found in fish.

To determine if a tainting evaluation is required as part of an EEM program, the following trigger should be applied (see also Figure 5.5):

"A mill need not conduct a fish tainting evaluation if at least one of the following conditions is met:

- fish contaminant analysis demonstrates that there is a potential health hazard; or
- there is no well-documented record of complaint [6] since 1992; or
- a sensory evaluation conducted in the previous EEM Cycle has demonstrated that tainting is no longer an issue.

Notwithstanding the above, each mill must conduct a fish tainting evaluation if fish contaminant analysis demonstrates that there is no potential health hazard and at least one of the following conditions is met:

---

[6] A well documented complaint can include (but need not be limited to) written complaints, fax or e-mail, documented telephone or verbal complaints to the mill or to government agencies.

- a sensory evaluation conducted in the last EEM Cycle has demonstrated that tainting is an issue; or
- there is a new record of complaint since the last EEM sensory evaluation.

The EEM study design for each mill should include a review of all pertinent information relative to tainting and an assessment of whether a tainting study is required in the current EEM Cycle.

Because the decision to investigate tainting associated with a pulp and paper mill is expected to occur on a site-specific basis, the local knowledge and expertise of regional authorities, such as Department of Fisheries and Oceans fisheries biologists or provincial ministry district biologists, is critical for ensuring that there is a sound basis for triggering tainting evaluations. The need for tainting studies should be based on a range of pertinent information including historical information, records of complaints, supporting information associated with a record of complaint, any previous tainting studies that were undertaken, and any subsequent environmental alterations or changes to mill operations that might be relevant.

Thus it is incumbent on the local regulatory authorities, taking into account the mill's operating requirements and the interests of all stakeholders, to assess whether a tainting study should be required as part of an EEM Cycle.

In cases where the responsible parties or the mill ownership do not have immediate knowledge of all prior information from which to assess whether tainting evaluations are required, some effort might be required to substantiate and/or validate existing information. In a similar manner, it might be necessary to confirm that the lack of complaints or documentation pertaining to a tainting problem is not a consequence of the abandonment of previously utilized fisheries resources by local users. A lack of utilization of fisheries resources should be due to tainting and not simply due to the perception that because a mill is in the vicinity it is not worth fishing there. At the same time, a lack of complaints should be due to a lack of tainting and not due to the fact that local users have not utilized the resource for many years because the fish were tainted at some time in the past. In most cases, such situations will be readily apparent, but there might be instances where further substantiation or validation is required.

When further substantiation or validation is required, a number of considerations may be necessary. Of primary importance is whether it is possible to assume that fish tainting is attributable to the discharge of mill effluent. Therefore, the site of the mill discharge, the location of natural and other anthropogenic discharges, physical characteristics (e.g. dilution capacity) of the receiving water body, and the location and utilization of fisheries resources near the mill should all be considered. Confirmation of the location and utilization of fisheries resources may require further investigation. For example, local libraries often have published histories of the region which could assist with understanding the availability and utilization of fisheries resources. Regulatory agency records could also contain relevant information.

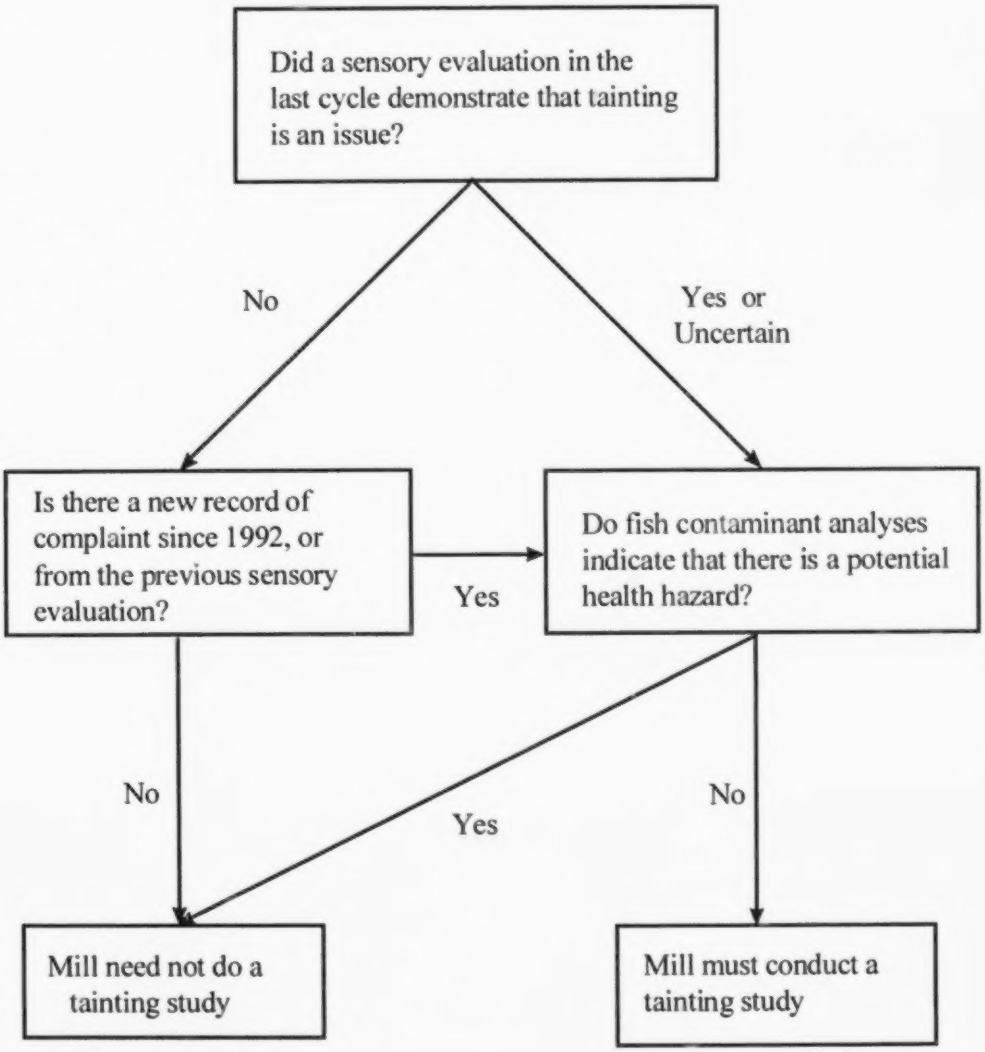


Figure 5.5 Tainting decision tree

Did a sensory evaluation in the last cycle demonstrate that tainting is an issue?
No
Yes or Uncertain
Is there a new record of complaint since 1992, or from the previous sensory evaluation?
Yes
Do fish contaminant analyses indicate that there is a potential health hazard?
No
Yes
No
Mill need not do a tainting study
Mill must conduct a tainting study

In cases where there is a history of complaints, the source and validity of such complaints should be reviewed. Creel records and/or surveys of local fishers can provide some guidance as to the basis of complaints. However, care should be taken to ensure that these records are unbiased and reproducible.

In cases where tainting evaluations have been undertaken in the past, the results and methods employed should be carefully reviewed.

5.3.2 Sample Collection and Sample Handling Methods

This section provides technical guidance on the collection and handling of tissue samples for assessing the usability of fisheries resources applying the difference from control sensory evaluation method.

The methods presented for collection and handling of samples generally refer specifically to finfish, however, similar methods can be applied for the collection and handling of shellfish for the assessment of taint.

5.3.2.1 Sample Collection

There are three principal methods for obtaining tissue samples, that have been exposed to pulp and paper mill effluent, for use in sensory evaluations. The methods, ranked in order of preference, are as follows:

1) collection of indigenous species;
2) *in-situ* or on-shore exposures; and
3) controlled laboratory exposures.

The most appropriate method to obtain fish for taint assessment will be dictated by site-specific conditions and the species of interest. Only if the requirements for the preferred method cannot be met should a mill proceed with the next-best method.

Collection of Indigenous Species

This is the preferred method of obtaining fish for sensory evaluations since the fish are likely exposed for extended time periods and the effects that ingestion of tainted prey may have on fish flavour (Mann 1951; Lovell and Sackey 1973) are also taken into account. These fish will also be most representative of the fish on which the complaints of fish tainting are based.

The fish that are used for tainting assessments should either be the species from which the complaints of tainting originated or a species that is commercially or recreationally important. NCASI (1987) provides additional criteria for the selection of appropriate species.

If the collection of indigenous species is to be used to obtain samples for sensory evaluations the following six requirements must be met:

1) *Both reference and exposure fish must be collected from the same waterbody and similar habitat types.*

It is common knowledge among fishers that fish from different waterbodies naturally taste different. For example, fish from lakes will generally have a different flavour than those from rivers. This is a factor of differences in diet, as well as water quality. Phytoplankton blooms and bacteria in lakes are known to affect fish flavour (Persson 1984; Martin and Suffet 1992). In the fish aquaculture industry the muddy/earthy flavour effects of blue green algae on fish have long been recognized (Thaysen 1936; Iredale and York 1976). Chemicals, such as, geosmin and 2-methylisoborneol occur ubiquitously in freshwater environments are well known for their effects on fish flavour (Lindsay and Heil 1992). Therefore, if fish were collected from a river downstream of a mill and compared to fish collected from a lake, any flavour difference found in sensory evaluations may be attributable to natural factors and not to the mill discharge.

2) *The same collection methods must be used in both reference and exposure areas. The order of preference would be trap netting, seining, angling, electrofishing, long lining and gill netting. However, only one method should be selected for the capture of all fish from each area. Fish must be removed from the water as soon as possible after capture.*

Similar to any of the other study components in an EEM program it is important to ensure that the same methods are applied to both reference and exposure samples. The primary uptake route of tainting compounds into fish tissues is through the gills (Persson 1984). Therefore, stress and respiration rate will greatly influence the rate of uptake. Consequently, if fish are captured in the exposure area by overnight gill net sets and the fish have been struggling in the nets for many hours before retrieval the level of taint will be elevated. In addition, natural compounds, such as lactic acid build-up in muscles, are known to affect fish flavour. For example, if fish were captured in the reference area by gill net and in the exposure area by trap net the added stress of the gill netting could result in the reference fish having a different flavour than the exposed fish.

3) *Fish in the reference area must be physically isolated (e.g. dam, falls, rapids, long distance) from the effluent-exposed population to be assured that there is no mixing of the two groups.*

The uptake rates of flavour inducing compounds can be very rapid, within minutes (Persson 1984) and can take days or weeks to depurate from fish tissues. Therefore, the data will be compromised if there is any chance that fish collected in a reference area were exposed to effluent within a couple of weeks prior to capture.

4) *The mill must be confident that fish collected in the exposure area have been exposed to effluent for a minimum of 48 hours.*

Studies focusing on tainting caused by pulp and paper effluents have demonstrated that fish flavour can be significantly affected within a few hours of exposure (BEAK 1993; NCASI 1987). These studies have also determined that a 48-hour exposure period is optimum in terms of the level of taint intensity and the cost-effectiveness of collections and exposures.

5) *Standard measurements of the fish collected for sensory evaluations, such as weight, length, sex and age, must be recorded.*

These are standard measurements that are recorded in any environmental study involving fish to account for data variability that may be attributed to any of these metrics.

6) *Six or more fish are recommended for pooling per sample to minimize the effects of natural flavour, colour and texture differences that may be attributable to individual fish; however, the minimum number acceptable for pooling is three fish per sample.*

Studies have shown that sensory evaluation panelists are able to distinguish natural flavour differences between individual reference fish. Also it is not expected that all fish will taint equally. Therefore, it is important to pool at least three fish for each sample to minimize these potential sources of variability. Furthermore, it is recommended that multiple samples be collected in reference and exposure areas to account for this natural variability.

Additional requirements that should be met if collecting indigenous species for sensory evaluations are as follows:

1) *Fish from reference and exposure areas should be approximately the same size and age and similar sex ratios should be collected from each area.*

It is common knowledge among fishers that older fish do not taste as good as younger ones. Metabolic rates are also different between older and younger fish and this would be expected to influence the level of taint in the tissues of exposed fish. Therefore, it is important to standardize, as closely as possible, the size and age of fish captured in reference and exposure areas. CPAR (1978) recommends that the maximum to minimum weight range should not exceed a factor of 1.2 in order to minimize potential differences in texture, colour and flavour. The ideal fish size would be similar to the size of fish from which the complaints of tainting originated. Furthermore, in the DFC test the larger the natural variability in fish flavour relative to the difference of interest (i.e. taint), the smaller the chances that a significant result will be obtained (Aust *et al.* 1985). Therefore, a sensory evaluation study should clearly indicate that all reasonable steps were taken to minimize the effects of natural variability.

2) *Only fish that are still alive at the time of collection should be used for testing.*

Fish decomposition, especially in warm waters, occurs quite rapidly and as a result fish flavour will be affected. It is important to only select live, healthy fish for sensory evaluations.

In-situ or On-Shore Exposures

If the above requirements for the collection of indigenous species cannot be met then the next best option would be to expose fish to effluent using *in-situ* caging or on-shore exposures. On-shore exposures are defined as holding vessels in which water is continuously pumped from the receiving environment through the vessels for the prescribed exposure period. If *in-situ* caging or on-shore exposures are selected the species from which the complaints of tainting originated should be used in the exposures. However, the appropriateness of this method and the species that is ultimately selected for exposure will be determined by site-specific conditions, hardiness of the species selected and the ability to meet the following requirements. Alternatively, surrogate species can be used provided it can be demonstrated that the surrogate species is expected to taint in a similar manner to the species on which the complaints of tainting are based. Previous studies have shown that pulp and paper effluent taint detection thresholds for rainbow trout were comparable to those for anadromous smelt that are known to have lipid levels as high as 15% (Beak 1996).

For this method to be applicable the following six requirements must be met:

1) *All fish used for the exposures must be collected from the same waterbody, preferably*

from a reference area within the mill's receiving environment. Alternatively, fish can be collected from another waterbody in the area or from a hatchery provided that all fish used for the study originate from the same waterbody.

The reasons for this requirement are the same as those presented for Requirement No. 1 in Section 5.3.2.1.

2) *All fish must be collected with the same collection method. The order of preference would be trap netting, seining, angling, electrofishing, long lining and gill netting.*

Only one method must be selected for the capture of all fish since the fishing method, depending on the level of stress induced, can affect fish flavour (e.g. lactic acid build up in tissues). A fish collection method should be selected that minimizes stress to the fish. The physiological state of the fish may affect the rate of uptake of tainting compounds (Persson 1984) therefore, only healthy, minimally stressed fish should be selected for exposure.

3) Fish must be exposed to effluent for a minimum of 48 hours.

Rationale discussed in Section 5.3.2.1, Requirement No. 4.

4) Standard fish measurements, such as weight, length, sex and age must be recorded

Rationale discussed in Section 5.3.2.1, Requirement No. 5.

5) Standard supporting measurements, such as, dissolved oxygen, temperature, conductivity, pH and water velocity must be recorded at each in-situ cage site or in on-shore exposure vessels.

All of these factors can influence the tainting response of fish when exposed to effluent. Conductivity measurements can be used to confirm effluent exposure concentrations. Other parameters, such as temperature and water velocity should be as similar as possible among cage sites.

6) Six or more fish are recommended for pooling per sample to minimize the effects of natural flavour, colour and texture differences that may be attributable to individual fish; however, the minimum number acceptable for pooling is three fish per sample.

Rationale discussed in Section 5.3.2.1, Requirement No. 6.

Further requirements that should be considered if this method of exposure is selected are as follows:

1) *All fish should be similar in size and age and the sex ratio should be the same at upstream and downstream locations.*

2) *In-situ cages or on-shore exposure tanks should be placed in areas where indigenous fish are suspected of being tainted.*

Exposures should be performed outside the Region of Discharge Generated Dilution (RDGD) similar to other components of the EEM program (Environment Canada 1997c). If in doubt of where indigenous fish may be tainted study designs may consider multiple exposure areas representing different effluent concentrations.

3) *Exposures should be conducted at the time of year when complaints of fish tainting are most common.*

There are generally seasonal changes in a mill's effluent quality, as well as seasonal changes in receiving water quality due to natural factors, such as surface runoff, algal blooms etc. Since exposure time is limited with this method it is important to perform the fish exposures at the time of year when fish tainting is most prevalent. This will also help to determine if natural factors are confounding or responsible for the flavour effects.

4) *Fish can be acclimated in dilution water for up to ten days, if practical, given the site-specific conditions.*

Fish exposures may be undertaken with fish that have not been acclimated provided reference and exposure fish receive the same handling and are obtained from the same waterbody. Previous studies have shown that acclimation periods from less than 24 hours to 10 days had no noticeable effect of the level of taint in fish tissues (Beak 1996).

5) *If the species of interest is unavailable or will not survive in captivity, surrogate species such as rainbow trout may be substituted.*

Rainbow trout have proven to be a good surrogate species for taint evaluations (CPAR 1978; BEAK 1994a,b, 1996).

6) *If in-situ cages are placed in fast-flowing environments the cages should be constructed in a manner that will provide the caged fish with shelter from the current. Cages should also be placed in similar water velocities in both reference and exposure areas.*

As mentioned previously, stress and increased respiration rate are expected to increase the level of taint in the exposed fish, as well as the build up of natural compounds that may alter fish flavour. Therefore, it is important to place cages in similar water velocities and these velocities should be measured and documented in the final report. The cages should also allow the fish to seek shelter from the current. This could be accomplished by ensuring that the upstream side of the cage is constructed of a solid material. Limited additional guidance, such as cage size, is provided in ASTM (1989).

Controlled Laboratory Exposures

The third fish exposure option would be controlled exposure in a laboratory. This option should be selected when site-specific conditions would prohibit the use of indigenous species or *in situ*/on-shore exposures.

Preferably, the species from which the complaints of tainting originated would be used in laboratory exposures provided they have no prior effluent exposure. Capture and handling of these fish should follow the procedures described above for exposures with *in-situ* cages or on-shore exposures. Alternatively, rainbow trout are recommended as the surrogate species to be used in laboratory exposures. Trout have been demonstrated to rapidly accumulate tainting compounds which can be readily distinguished by sensory panels because of the mild inherent flavour characteristic of the fish flesh. These fish are available from hatcheries across Canada and therefore, could be used as a standard species for laboratory tainting evaluations.

Test solution replacement in the laboratory assay can be either continuous flow (with sufficient flow-through to maintain 70% saturation of oxygen in the test solution) or static renewal (with solution replacement every 24 hours). Studies have shown that there is no significant difference in taint detection thresholds between static and flow-through exposures (Beak 1996, CPAR 1978).

Twenty-four hour composite samples of the effluent are recommended for testing although grab samples would be acceptable for site-specific situations. Fresh samples should be collected each day and transported to the laboratory in clean non-toxic containers (e.g. polyethylene or polypropylene) that are filled to the top so there is minimum headspace. Transportation and storage conditions should be the same as for effluent samples collected for routine toxicity tests. Because many of the tainting compounds are volatile and also biodegradable, fish exposures must be initiated within 24 hours of sample collection.

The dilution water for the laboratory exposures (both effluent and control exposures) should be the receiving water from the reference site for the mill, where it is practical (i.e. on-site laboratory exposures). Otherwise, the laboratory culture water could be used. However, comparison with receiving waters is recommended.

For this exposure method to be applicable the following four requirements must be met:

1) *Fish must be exposed to effluent for a minimum of 48 hours.*
2) *The number of fish per exposure must conform with the sample size (# of grams) required for the sensory evaluation and a minimum of three fish will be exposed to each effluent concentration tested.*
3) *Aeration in each test vessel must be maintained at a constant level.*

Studies have shown that tainting compounds can be stripped from the effluent with aeration. Therefore, it is imperative that the minimum rate of aeration to maintain 70% oxygen saturation be used and that all exposure vessels receive the same aeration rate. The aeration rate for each vessel must be recorded and must not exceed 7.5 mL/L/min.

4) *Conductivity must be recorded in each vessel to confirm nominal effluent concentrations and in flow-through systems conductivity must be recorded frequently to ensure that diluter systems are continuously providing the prescribed effluent concentrations.*

Dissolved sodium measurements can also be made to confirm effluent concentrations. Additional requirements that should be met when using laboratory exposures are as follows:

1) *All fish should be within the same size range and age.*
2) *Loading density of fish in each test concentration should be minimized and must not exceed a maximum of 10 g/L/day.*

The loading rate in each exposure vessel must be reported.

3) *Fish used in laboratory exposures should have undergone a period of at least 10 days acclimation and depuration prior to the test.*

This may not be practical if working with wild species that are not easily maintained in captivity. However, an acclimation period is recommended if working with hatchery raised species so that flavour effects from water quality conditions at the hatchery can be depurated.

4) *Fish should not be fed for 48 hours prior to the test in order to minimize off-flavours*

that may result in static systems from fouling of the water.

5) *Lighting should be maintained under the standard conditions recommended for fish culturing and aquatic toxicity testing.*
6) *Test solutions should be maintained within ± 2°C of the acclimation temperature.*
7) *The effluent concentrations selected for exposures should bracket the range of concentration in the receiving environment where tainting is suspected to be occurring.*

5.3.2.2 Sample Handling

The goal of these methods is to immediately stabilize harvested fish in order to prevent any changes due to quality deterioration which could interfere with the evaluation of flavours from exposure to effluent. The important factors are cleaning, protecting from exposure to air, and chilling (first by holding on ice, then by freezing as soon as possible). Under no circumstances should the viscera come into contact with the flesh that will be used for sensory evaluations.

The following standard operating procedures must be followed:

- Fish must be cleaned by eviscerating and removing kidney and gills immediately after capture.
- The carcasses must be rinsed in clean water, such as distilled water or clean water from the reference area and immediately placed on ice with the body cavity facing downward to allow for drainage. A sufficient volume of ice should be used so that none of the fish come in contact with the drainage. Fish taken from laboratory exposures should also be rinsed in clean water prior to freezing.
- The cleaned fish should be placed in a freezer within eight hours of collection and maintained at least at -20°C until sensory evaluations are conducted. The lower the temperature of frozen storage, the more stable the samples will be.
- The fish should be tightly wrapped in foil or plastic wrap to shield from air and placed in double zip-lock bags with as much of the air removed as possible. Ideally, the fish samples should be vacuumed sealed prior to freezing.
- Samples must remain frozen at all times during shipping from the field to the sensory evaluation laboratory. This can be attained by shipment with dry ice or by refrigerated transport truck.
- Samples should not be frozen for more than one month before testing.

5.3.3 Sample Preparation and Sensory Evaluation Methods

5.3.3.1 Background

In most studies investigating the effects of pulp and paper mill effluents on fish flavour the sensory evaluation method has generally been difference testing using the triangle test method. In this method a panelist is presented with three coded samples, two which are the same, and is asked to identify the odd sample. The main problems in using this method for the EEM program are as follows:

- the procedure is used to determine whether any perceptible sensory difference exists between samples of two products and any difference, not only the presence of mill effluent taint will be used by the assessor in examining the samples (e.g. colour, texture);
- the difference may be in a single sensory attribute or in several;
- the test is applicable when the nature of the difference is unknown, but it does not determine the magnitude or direction of difference and without panel training it does not identify the attribute responsible for the difference;
- it is applicable only if the products are homogeneous, and does not account for natural variations in the product, such as those found when working with fish;
- since the test simply identifies a difference and not the magnitude of the difference the data does not allow for a national comparison of pulp and paper mills across Canada to determine if the PPERs are adequate to prevent fish tainting.

Preference and acceptance testing which are also sometimes used to test for the acceptance of a product are subjective tests which determine consumer preferences for and attitudes towards different products. They are designed to be used with large numbers of untrained assessors (generally around 100) and measure opinions and biases rather than product characteristics. In this type of testing, there is no right or wrong answer, as any like/dislike response to the sample is valid. This creates problems in the interpretation of data regarding the quality of the samples (i.e. the presence or absence of taint) and in the final assessment of the adequacy of the PPERs to prevent fish tainting. Also affecting the interpretation of the data is the fact that it is not uncommon for an assessor to prefer the taste of an effluent exposed fish, however, when cognisant of the fact that the difference in taste is attributable to exposure to pulp and paper effluent the taste is no longer acceptable. A prime example of the fallacies of using this type of testing for measurement of quality in fish is given by Kosmark (1987).

The difference from control (DFC) test, which is recommended for the EEM program for the evaluation of taint in fish flesh, is often used in quality control situations for the measurement of a difference and the estimation of the size of any differences found, taking into account the natural variability in the product.

The advantage of this method over tests, such as the triangle test for EEM monitoring is that the nature of the difference required is easier to define (i.e. taint) and the assessor is less influenced by other factors in the sample. The data gathered can be analyzed through standard statistical tests to give measurements of intensity of taint relative to the control samples.

Fish are well known to have naturally variable flavours. For example, statements are often made by fishers that older fish do not taste as good as younger ones. The main advantage of the DFC test in the context of the EEM program is that this test recognizes that most foods, being natural products, have some variance in flavour (Yantis 1992). This natural variance is taken into account in the experimental design and statistical analysis (Section 5.3.4). Comparison studies of sensory evaluation methods (i.e. triangle versus DFC tests) have shown that the use of the triangle test with a highly variable product, such as fish, results in a high frequency of false statistically significant results compared to the DFC tests which measures perceived difference in taste due to other than product variability (Aust *et al.* 1985).

The results from the DFC tests once entered into a national database will also allow the pulp and paper industry to look for relationships between mill processing methods and the degree of taint in fish.

Difference From Control (DFC) Test Method

In the Difference-from-Control Test, samples of treated fish which have been exposed to pulp and paper mill effluent either under field or laboratory conditions and blind (hidden) control samples are compared with an identified control sample. The assessment of the presence of a difference based on a specified characteristic (i.e. taint) and the degree of that difference are made using a standard sensory evaluation ballot. Further details of this test method are given in Meilgaard *et al.* (1991) and Munoz *et al.* (1992).

It is preferable to use a ballot with an unstructured line scale because it helps to minimize central tendency error since in numbered/category scales panelists have a tendency not to use the lowest or highest three numbers in the series (Meilgaard *et al.* 1991). A new ballot should be used for each set of samples which will help to reduce errors due to patterns effects. If a panellist anticipates that there is some sort of pattern they will tend to compare the scores for the present sample to the scores that they assigned to previous samples. Figure 5.6 provides an example of a standard test ballot for the DFC test. The responses from the unstructured line scale are decoded into numerical data to the nearest one tenth of a centimetre (e.g. 9.3). Data collection for each replicate sample includes the recording of the values for each of the samples given by each of the panelists so that individual assessor's data can be tracked and the panel performance and consistency assessed across replications.

It is recommended that when the panelists are presented with a coded sample they should also be given a freshly cooked identified control sample with which to compare the odour and flavour of the coded sample.

Section 5.3.4 provides technical guidance on experimental design and statistical analysis for the DFC test method used in the context of an EEM program. The following subsections provide technical guidance on facilities, sample preparation, selection of panelists, training and monitoring, and sample presentation methods for the DFC test.

5.3.3.2 Sample Preparation

The goal of these methods is to prepare samples for serving to a sensory panel without causing quality changes during the handling of raw fish, flavour changes due to cooking and changes in taint levels due to loss of volatiles by subsampling cooked fish. The methods also minimize any differences among samples, such as texture or colour, that will bias the panelists' decisions (i.e. stimulus error). The sensory evaluation laboratory must have detailed standard operating procedures (SOPs) for the preparation of samples and these SOPs should be reviewed by the technical advisory committee for each mill. The SOPs should encompass the following points and those recommended by ASTM (1996).

- Thawing of fish is done at 4 °C only until thawed enough to handle. Generally, fish samples are removed from the freezer the night prior to sensory evaluation. Reference and exposure fish should be thawed in separate refrigerators.
- Generally, skinless fillets are used for testing. When filleting and skinning, care should be taken to keep the flesh away from any slime, scales, etc. from the skin and to keep the fillets chilled during handling and processing. All utensils in contact with the samples should be thoroughly cleaned between samples following SOPs to prevent cross-contamination of tainting compounds.

Figure 5.6 Example of a sensory evaluation ballot used for DFC tests with fish exposed to pulp and paper mill effluent. (D. Farara, pers. comm.)

Name:______________________________ Date: __________________

You have been given a control sample and a coded sample. Proceed as follows:

1. Eat a piece of cracker and rinse your mouth out thoroughly with water before beginning. Expectorate water in the container provided.
2. Begin with the control sample, tear open package and take three short sniffs. Note the odour present. Evaluate the odour of the coded sample following the same procedure. Return to the control sample and divide the sample in half using a fork. Taste and expectorate the sample. Rinse your mouth thoroughly with water and expectorate. Evaluate the taste of the coded sample using the same procedure.
3. Place a vertical line across the horizontal line at the point that best describes the degree to which the coded sample differs from the control **in terms of odour and flavor.**

Note: *Remember that a control sample may be included as a coded sample some of the time.*

Code No: ______

000

No difference/
Same as Control

Extremely
different from
Control

Comments:

Note: OOO denotes the reference sample upon which panelists were trained. This sample represents a high degree of taint for that particular sample in question.

- Blend fillets to create homogeneous samples for each treatment/replication. Composite samples are necessary to reduce the effect of individual fish variability in uptake of effluent into the fish, as well as natural flavour, texture and colour variability. A minimum of three fish (i.e. six fillets) must be homogenized for each sample. In order to minimize texture differences the same number of fillets for reference and exposure samples must be blended together. Detailed SOPs that minimize texture differences between reference and exposure fish must be documented. For example, blending of fresh rainbow trout fillets and subsequent freezing of the blended flesh results in a totally different textured sample than blended fillets that were previously frozen. Panelists generally prefer the texture of the latter method (D. Farara, pers comm.).
- Individual samples are wrapped in heavy aluminium foil and coded with unique (for each treatment/replication) three-digit codes. This preparation format allows the samples to be completely sealed for cooking and does not require subsampling after cooking. This means that the assessor is able to evaluate all the volatiles present in the fish sample.
- The individual samples should contain from 7 to 15 g of fish, depending on the availability of fish and number of samples required. The minimum sample size should of sufficient size to allow a panelist to taste the sample at least twice.
- Samples are cooked in a steamer for seven to ten minutes depending on the number of samples being cooked. The cooked packets are presented directly to the panel in a random balanced serving order (i.e. panelists evaluate the samples in a different order) and are not unwrapped during or after this preparation step. Steaming prevents the addition of flavours from the cooking process, such as those due to browning. Reference and exposure fish must be cooked in separate steamers.
- Samples should be presented to the panelists under red lighting to minimize stimulus error due to colour differences. There is a high probability that there will be colour differences between samples, especially when dealing with salmonid species and these differences will need to be masked so the panelists' scores are unbiased.

5.3.3.3 Selection, Training and Monitoring of a Sensory Evaluation Panel

The purpose of this section is to describe the requirements for selection, training and monitoring of panelists that will be used to measure the presence and intensity of taint in fish exposed to pulp and paper mill effluent in either field or laboratory conditions for an EEM program. For the DFC test the panelists must be screened and trained. The following information describes standard methods for the selection, training, monitoring and use of sensory panels. All of these sensory evaluation methods have test protocol requirements and potential sources of error which make it essential that they be conducted by or supervised by a person with specific experience in sensory analysis, especially in panel selection, training and monitoring. The sensory evaluation supervisor must have sufficient experience to be able to recognize potential sources of error. The qualifications of the sensory evaluation supervisor and technicians should be documented in the final report.

Facilities

All sensory tests and screening of panelists must be conducted in an appropriate sensory testing facility which meets the physical requirement guidelines documented in ASTM (1986). The sensory evaluation facility must include the following:

- individual booths for panelists;
- temperature control to 22 °C ± 2 °C;
- adequate and appropriate lighting for the task, generally 750 to 1000 lux in each booth;
- proper ventilation which ensures removal of odours from the testing area;
- coloured lighting to mask colour differences between samples;
- separate rooms for testing and preparation of samples; and
- control of noise and distractions during testing.

Selection of Potential Panelists

Selection of assessors who have the basic sensory acuity to perform the tests needed is essential to the collection of scientifically defensible data. The personnel selected for participation in sensory evaluation panels must be trained and validated in the collection of objective sensory data for the presence and intensity of taint in fish using the DFC test. This includes training in the use of the DFC test methods and instructions on the use of the ballots, sample handling during assessments, rinsing procedures, etc. They are also given experience with known samples of tainted and control fish, and their ability to participate in the test is further assessed through these samples.

For the DFC test method the panel is generally composed of 12 or more trained panelists. For the EEM program the panel size must not be less than 10 members. Generally, to achieve a panel size of 12 or more members it will necessary to screen at least three times that number (i.e. more then 36 candidates).

Pregnancy is the only factor that will automatically eliminate a candidate, as may other health factors (e.g. allergies, sinus problems, some medications) that could interfere with sensory acuity. Excluding pregnant individuals is done because pregnancy is known to affect the senses of taste and smell (Poste *et al.* 1991).

The first step in panel selection is a personal suitability interview. This is generally conducted during the recruitment phase of potential candidates. During the personal suitability interview, the candidate is assessed for such characteristics as:

- interest in panel work and in the particular study
- availability and promptness
- cooperativeness, such as a willingness to refrain from using any perfumes or other scented lotions or refrain from eating, smoking or drinking at least one hour before sensory evaluations, as these factors may interfere with the individual's performance or will be a distracting odour to other members of the sensory panel
- general health (chronic health problems which could interfere with sensory panel performance: for example, allergies, chronic nasal problems, medications, etc.)

Generally, a sensory evaluation questionnaire is developed which asks these questions, as well as others such as, age, sex and their liking of fish. There is no limitation on the age of panelists, or whether or not they smoke as the ability to perform the test is assessed during the screening phase. However, it is strongly recommended that panelists be at least 16 years of age.

Knowing the personal qualities of the candidates who have been selected to participate in the study will help to ensure that they will be able to conduct the sensory assessments in a consistent and professional manner.

Screening of Potential Panelists

The purpose of these procedures is to select panelists with the sensory abilities to perceive the sensory attributes for which the samples are being examined (i.e. taint caused by pulp and paper mill effluent) and with the personal qualities required to participate in the tests.

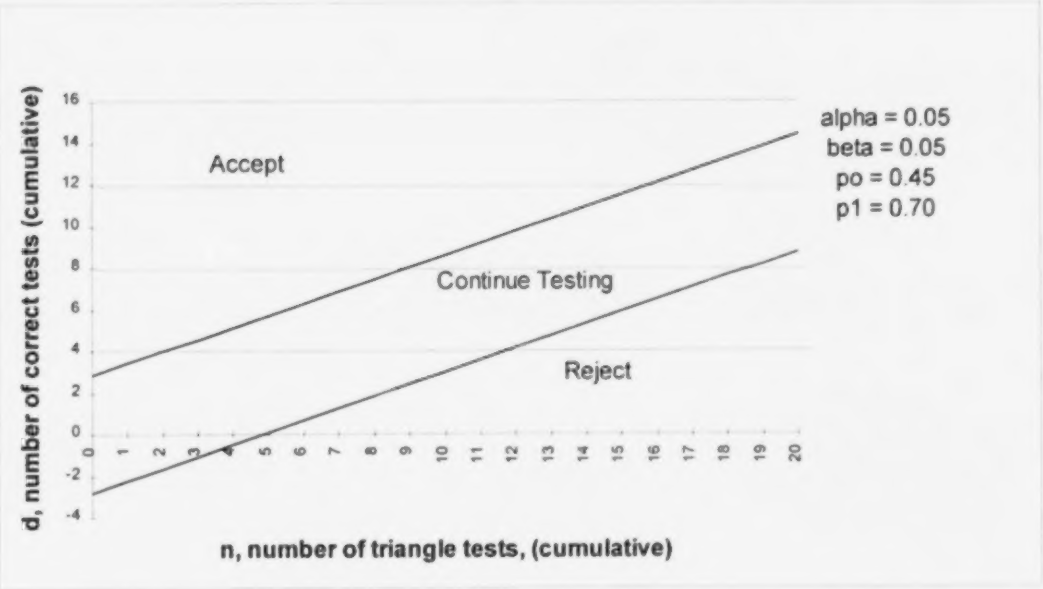
The panelists should be screened for their taste acuity following the "sequential testing" method. The group of tests which fall under the heading of "sequential testing" are actually difference tests which are used with pre-set statistical criteria. They allow the decision of "accept for panel", "reject for panel" or "continue testing" to be made through the use of a pre-set graphical representation of the statistical criteria (Figure 5.7). These tests have been described in the sensory literature (Amerine *et. al.* 1965; Meilgaard *et al.* 1991; and Munoz *et al.* 1992) as being meant to "economize" on the number of tests which must be conducted in order to accept or reject a panel trainee while providing a statistical basis on which to base this decision. This method is based on the *a priori* selection of the statistical levels of α and β probabilities for the level of performance of the assessors and the use of difference tests, such as, triangle, duo-trio, and paired comparison-difference. For the EEM program the values for α and β in sequential testing must be established at 0.05.

The selection of appropriate samples for screening of panelists is a critical step in the overall sensitivity of the DFC method. Samples selected for screening must not be too easy (i.e. the taint is easily detected by most panelists) but at the same time should not be too difficult (i.e. most of the panelists cannot detect any taint). While the effluent exposure concentrations will be set on a mill by mill basis, they should represent relatively small differences so that the sensory laboratory will be able to clearly demonstrate that selected panel has the ability needed to collect useful data. As a general rule, fish exposed to effluent concentrations that represent concentrations in the river where fish tainting is suspected to be occurring, and fish exposed to five and ten times this concentration should be used for panel screening. To obtain these samples fish will have to be exposed to known concentrations of effluent, which will generally require laboratory exposures following the methods presented in Section 5.3.2.1. A graphical example of sequential testing using three effluent concentrations and the forced-choice triangle sensory test is provided in Figure 5.8. For this particular example the group taint detection threshold was around 1% pulp and paper effluent. Panelists who correctly identified the effluent exposed sample enough times to fall within the accept category for all three effluent concentrations were selected for further training.

The following are some of the critical points that are common to all of the sensory test methods used in an EEM program:

- The time of day when testing is conducted can influence the results and it is agreed among sensory specialists that late morning and mid-afternoon are generally the best times for testing (Poste *et al.* 1991).
- Prior to entering the panel room, the panelists should wash their hands with odourless soap and rinse to remove any trace of residue. During panel sessions, panelists should clean hands of any fluid residue from the samples using a paper napkin dipped in odour-free water in order to prevent cross-contamination of samples and physical carryover of any stimuli, especially odours.
- Oral rinsing between samples using neutral materials is standard practice during sensory evaluation sessions. The purpose of rinsing is to prevent the carry-over of stimuli from one sample to the next. Odour-free water, heated to 50 °C (e.g. bottled or distilled) is recommended along with unsalted soda crackers. When crackers are used, they must be followed by a water rinse to remove any residue from the mouth. It has been found that some bottled waters have an odour and taste, therefore, it is critical that the sensory evaluation supervisor confirm the quality of each batch of water before use. For the EEM program diluted concentrations of lemon juice must not be used as it can interfere with low levels of sourness in the samples which is often a

Figure 5.7 Graph for sequential analysis for selection panel trainees using triangle tests



flavour characteristic of fish exposed to pulp and paper effluents.

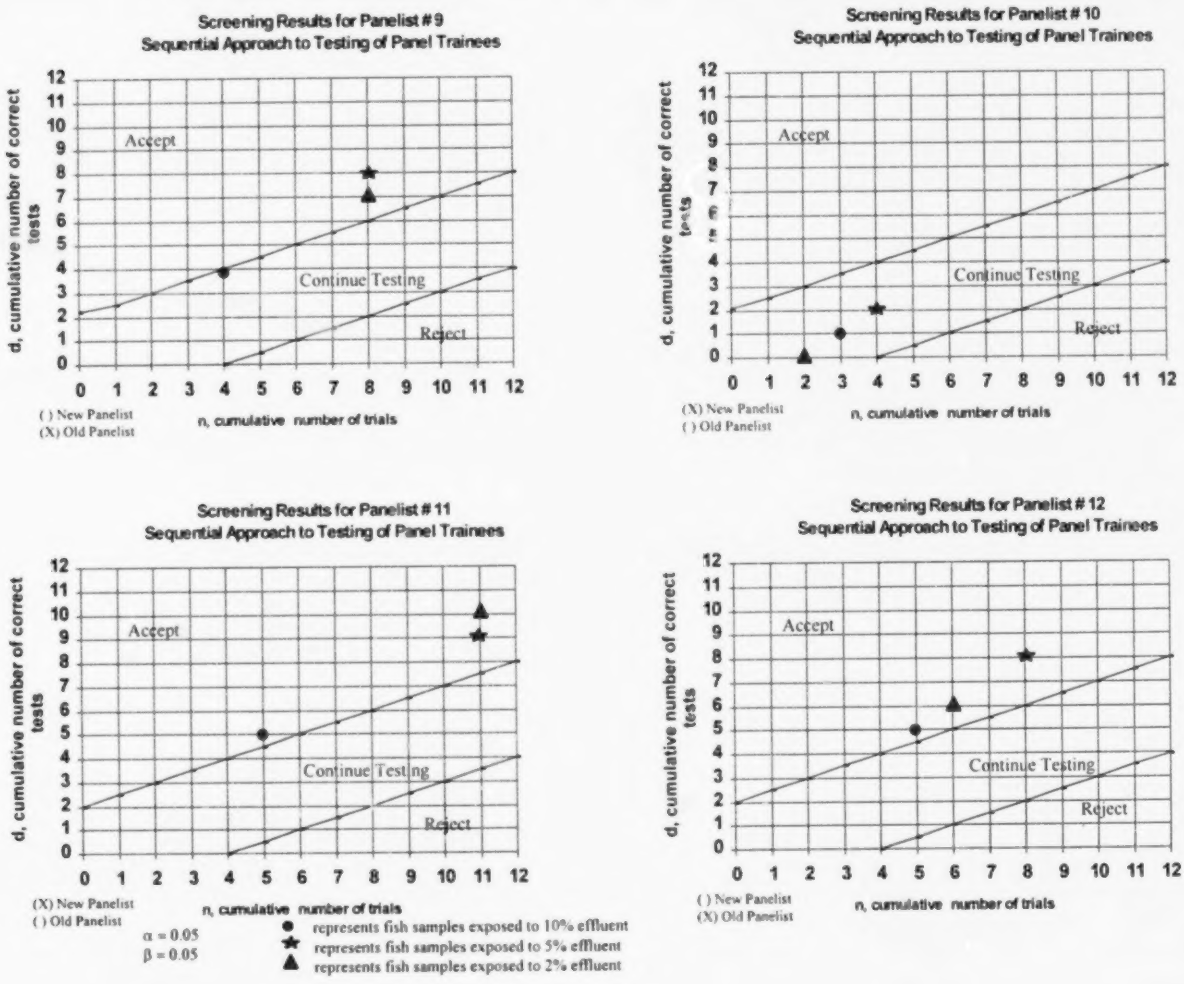
Time between sets of samples is required to avoid carry-over of odour/flavours from the previous sample. Based on the feedback of panelists from previous studies a minimum of ten minutes should be taken between each set of samples and no more than five sets of samples should be served within a single seating or session (BEAK 1991, 1992, 1993, 19941,b, 1995, and 1996). Panelist fatigue is minimized by rinsing and also by limiting the number of samples which are analyzed at each session. The time taken between individual samples within a set of samples (i.e. the known control and the coded sample) should be 10 to 20 seconds (ASTM 1985).

- The panelists' instructions for handling samples during evaluation sessions that should be on the ballots include:
  - open sample packets by tearing off one folded end and down one side of the foil;
  - evaluate the odour of the samples with three short sniffs first then close the packet temporarily;
  - to taste, divide the sample in half and do the assessment. Use second half as needed for re-tasting.
- Panel methods for the analysis of taint includes the use of "warm-up" samples at the beginning of each test session which would include samples with a known level of taint that the panelists can recognize. This allows the sensory systems to "re-experience" the sensory attributes which will be evaluated before the actual test samples are begun.

Training of Panelists

This step is essential to establish a calibrated panel and should be expected to take between two to three weeks of daily panel sessions. However, once the panel is trained, subsequent testing for a different mill or using a different fish species will require at least one week of additional training with the new effluent and/or species which acts as a "refresher" and re-calibration for the next data collection panels. Panelists are trained to recognize the control (depending on the species of fish used) and to recognize the taste of pulp and paper tainting compounds in the fish species being studied. Proper training of panelists reduces the probability of making a Type II error (i.e. concluding there is no taint when taint really exists) by improving an individual's ability to recognize and identify the taint and improve their sensitivity and memory so that they provide precise consistent measurements (ASTM 1981).

Figure 5.8 An example of the sequential approach to panel selection using fish exposed to three concentrations of pulp and paper mill effluent (from Beak 1996)


Screening Results for Panelist # 9
Sequential Approach to Testing of Panel Trainees

Screening Results for Panelist # 10
Sequential Approach to Testing of Panel Trainees

Screening Results for Panelist # 11
Sequential Approach to Testing of Panel Trainees

Screening Results for Panelist # 12
Sequential Approach to Testing of Panel Trainees

The training program includes several sessions during which the assessors become trained to:

1) use the correct procedures for the DFC test method including:
    - sample handling;
    - evaluation techniques to reduce fatigue (e.g. rinsing time);
    - analytical approach to testing and disregarding personal preferences;
    - correct usage of the ballot for recording data;

2) accurately evaluate the actual samples through the presentation of known, prepared samples which are representative of concentrations which occur in the receiving environment where tainting is suspected to be occurring and concentrations that are five and ten times that level:

    - begin with large, easily perceived differences, such as ten times the effluent concentration in the receiving environment where fish tainting is suspected to be occurring;
    - progress to smaller but easily perceived differences;
    - continue until assessors can demonstrate successful "calibration" (data analysis here can be ANOVA comparing treatments and panelists as main effects and use the treatment x panelist interaction to measure the overall performance of the panel);
    - panelists with scores that frequently fall outside two standard deviations of the panel mean should be removed from the panel.

Training calibrates the panel and provides the individual assessors, the sensory evaluation supervisor, the mills and the regulatory authorities with confidence in the panelist's sensory abilities required to perform the test. The panelists that are consistent in their scoring of samples and who score samples within the expected range are the ones who should be selected for the sensory panel.

Reference Samples

The use of reference samples or standards is critical in both the training of panelists and calibrating of panelists during testing (Malcolmson 1995). In previous DFC testing with fish exposed to pulp and paper effluent the panelists had difficulty determining what intensity of taint would warrant the highest score (D. Farara, pers. comm.). In order to anchor the endpoint at the upper end of the DFC line scale fish samples were tainted to a level that was very obvious to the panel. During the training period the panelists were trained that this level of taint intensity would score a 13 on a 15-point line scale (Figure

5.4, labelled as "000"). This standard reference sample is now used for the warm-up sample at the beginning of a session and as a hidden standard for panel management and QA/QC when monitoring the reliability of individual panelist's results.

In the context of the pulp and paper EEM program all labs must establish such a reference standard so that the data from different sensory evaluation labs will be comparable once the data is entered into a national database. It is important that the selected reference clearly represents the taint and that it does not change over the course of the study (Malcolmson 1995).

Panel Management (QA/QC)

Quality control of the sensory panel includes the tracking of individual panel results to assess any changes in sensory acuity during the course of a study. If this occurs, retraining is necessary. This can be assessed through data analysis, such as, ANOVA, where the main effects become: Samples, Panelists, Replications. The inclusion of panelists in this allows the determination of the consistency with which the individuals rate the samples, as well as their consistency across replications. Simple graphical presentations showing the scores of individual panelists are also helpful in identifying panelists who are not performing as expected, based on the trends seen with the panel as a whole.

Part of the experimental test design for assessment of pulp and paper mill effluents includes the addition of hidden control samples and hidden reference standards to the test samples. Hidden control and reference standard samples also allow monitoring of the panel during test sessions. Scores given to the blind control sample provide a check of internal validity of the panel and serves as a measure of the effect of asking the difference question (i.e. placebo effect)(Munoz *et al.* 1992). The insertion of hidden control and standard samples also helps to minimize the error associated with pattern effects if the panelists learn which samples to expect in each session.

It may be necessary to replace panel members, if any withdraw from the panel. In this case, it will be necessary to repeat the selection procedures and the initial training panels for the new member(s). Also if a particular panelist shows significant changes in their sensory acuity during the study it may be necessary to drop this panelist's results from the data set before statistical analysis. Alternatively, similar to the data analysis and interpretation guidance provided for other components of the EEM program (e.g. benthic invertebrates, adult fish survey) the data can be statistically analysed with and without the outliers. Substitution of panel members during the same study should be avoided (ASTM 1981).

5.3.4 Experimental Design and Statistical Analysis

The experimental design selected for a study will influence the accuracy of the data (Malcolmson 1995). The following section provides a recommended design for an EEM study using exposed fish from either *in-situ* caging/on-shore exposures or collection in the receiving environment. The design takes into account the natural variability in fish flavour that may be expected and the variability in the uptake rates of tainting compounds by individual fish that may occur in the exposure area.

The null hypothesis in the DFC test for EEM studies is that the reference area equals the exposure area, with respect to the mean panel scores.

The procedure follows Yantis (1992), except that two "exposed" samples will be included, as opposed to just one, and testing is performed in two or more sessions. Thus, there will be two reference samples and two exposed samples. Each sample will be a composite of at least three fish and preferably six or more from the same station within an area (reference or exposed), either from the same cage if cages are used, or from the same area of capture if wild caught fish are used. In other words, fish should be collected from two stations located in a reference area and two stations located in the exposure area, resulting in four groups of fish. Each of these samples will be scored "blind". The two stations in the exposure area must be located in the same effluent concentration.

At one of the reference stations, a "known reference/control" sample will be produced also. This known reference sample will come from the same fish composite as one of the "blind reference" samples for that station. The "known reference" material will be presented with each of the "blind" samples. The flavour of the blind samples will be scored in comparison to the known reference sample. A "blind standard" sample (a positive control from a large batch of exposed fish that has been previously tested) will also be presented for scoring in at least one session as part of the panel management and QA/QC.

Each of at least ten panelists will be presented with four blind samples for scoring (excluding the blind standard), in at least two sessions. They will be asked to score each blind sample on a 0 to 15 scale where 0 is no difference from the known reference and 15 is extremely different. An example of complete data set will follow the structure of Table 5.14.

The data will be analyzed by Analysis of Variance (ANOVA), using a randomized complete block design, with panelists and sessions as blocks. An example of the resulting

ANOVA table using the minimum number of panelists (i.e. 10) is provided in Table 5.15, assuming a fixed effects model.

The degrees of freedom are calculated as follows:

- Total = number of scores (N) - 1;
- Session = number of sessions - 1;
- Panelist = number of panelists - 1;
- Reference vs Exposure = number of areas - 1; and
- Pure Error = number of panelists x sessions x areas.

The fixed effects model does not permit panelists or session dates to be viewed as representing a larger population of panelists or session dates. Therefore, conclusions will be valid only for the particular panelists and sessions in the study. If panelists and sessions are viewed as random effects, while areas are fixed effects, then a mixed model is implied. In this model the residual error would be partitioned into component interactions and these would be used to derive appropriate error terms (as opposed to pure error) for testing each main effect. Snedecor and Cochran (1971) provides additional details on the mixed effects model.

The "pure error" variance (mean square) represents the variability between stations (i.e. cages or collection areas) within areas (reference or exposure). It differs from the "pure error" of Yantis (1992) in that it is pooled across areas. This is appropriate, since the hypothesis pertains to areas and since there is natural variability within both reference and exposure areas, but homogeneity of variance between areas should be examined. If the two variances are significantly different, then transformations may be used if possible to stabilize the variance, before proceeding to compare the means. Otherwise, the difference in variability between the two areas (reference and exposure) is the main finding of the study.

The effect of primary interest is the Reference vs. Exposure effect. This mean square is divided by the "pure error" mean square to obtain F. Standard tables of critical F are consulted to determine the significance of the effect. The panelist and session effects are also examined as a check on panel performance. Significant panelist effects are often observed and do not invalidate the test; however, if one panelist differs strongly from the others, he/she may be removed to improve panel performance. Significant session effects are not expected; if observed, they should be investigated to ascertain whether some unexpected factor may have influenced panel performance.

After the panel training period, power analysis should be performed to identify the magnitude of Reference vs. Exposure difference that can be detected in two sessions. If this detectable difference is considered to be too large, the number of sessions can be increased to increase the resolution of the study.

Table 5.14 Data from degree of difference test

Panelist	Session 1				Session 2			
	Reference 1*	Reference 2	Exposure 1	Exposure 2	Reference 1	Reference 2	Exposure 1	Exposure 2
1	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-
9	-	-	-	-	-	-	-	-
10	-	-	-	-	-	-	-	-
SUM	-	-	-	-	-	-	-	-

* Reference 1 is used as the known reference for each set of samples.

Table 5.15 Analysis of variance for Table 5.14 data

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	P
Total	79	-	-		
Session	1	-	-	-	-
Panelist	9	-	-	-	-
Reference vs. Exposure	**1**	-	-	-	-
Pure Error	**40**	-	-		
Residual	28	-	-		

5.3.5 Reporting Requirements

All sensory evaluations reports must contain enough detailed information to allow the RAO to judge the appropriateness of the methods and to evaluate the reliability of the results. The report should also provide sufficient data so that the study can be repeated.

The final sensory report should contain all of the following information:

- the objectives of the study;
- the experimental design;
- description of the testing facility, including a description of lighting, ventilation and

climate control;

- qualifications of the sensory evaluation supervisor;
- sample collection methods and all supporting field and laboratory measurements;
- detailed sample handling and preparation methods;
- methods and results of panel screening;
- methods and results of panel training;
- demographics of the sensory panel;
- QA/QC results from panel management that would demonstrate the reliability of the panel;
- methods and results of the DFC tests, including raw and summary data;
- detailed statistical methods and results including ANOVA tables; and
- results of *a posteriori* and *a priori* power analysis and identification of the effect size that could be reliably determined.

5.4 References

Alldredge, J.R. 1987. Sample size for monitoring of toxic chemical sites. Envir. Monitoring Assess. 9:143-154.

Amerine, M.A., R.M. Pangborn, and E.B. Roessler. 1965. Principles of Sensory Evaluation of Food. Academic Press, New York.

ASTM. 1981. Guidelines for the Selection and Training of Sensory Panel Members. ASTM STP-758. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1985. Standard Practice for Establishing Conditions for Laboratory Sensory Evaluation of Foods and Beverages. ASTM E480-84. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1986. Physical Requirement Guidelines for Sensory Evaluation Laboratories. ASTM STP 913 American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1989. Standard Practice for Evaluating an Effluent for Flavour Impairment to Fish Flesh. Report D 3696-89. Committee D-19. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1996. Standard Practice for Evaluating Effects of Contaminants on Odor and Taste of Exposed Fish. Report E 1810-96. American Society for Testing and Materials, Philadelphia, PA.

Aust, L.B., M.C. Gacula, Jr., S.A. Beard and R.W. Washam, II. 1985. Degree of difference test in sensory evaluation of heterogeneous product types. J. Food Sci. 50: 511-513.

Bartlett, J.R., P.F. Randerson, R. Williams and D.M. Ellis. 1984. The use of analysis of covariance in the back-calculation of growth in fish. J. Fish. Biol. 24: 201-213

Beak Consultants Limited (BEAK). 1991. Fish Tainting Potential of Eurocan Effluent in the Kitimat River, British Columbia.

Beak Consultants Limited (BEAK). 1992. Evaluation of the Tainting Propensity of Eurocan Final Effluent.

Beak Consultants Limited (BEAK). 1993. The Effect of Eurocan Effluent on Oolichan Flavour - 1993.

Beak Consultants Limited (BEAK). 1994a. Kitimat River Eulachon Fishery and the Effects of Eurocan Effluent on Eulachon Flavour and Egg Survival - 1994.

Beak Consultants Limited (BEAK). 1994b. Comparison of Laboratory and In-Situ Detection Thresholds - September 1994. Report for Eurocan Pulp and Paper Co., Kitimat, B.C.

Beak Consultants Limited (BEAK). 1995. Eulachon Flavour Impairment, Egg and Fry Survival, and Fishery Assessment Studies - 1995. Report for Eurocan Pulp and Paper Co., Kitimat, B.C.

Beak Consultants Limited (BEAK). 1996. Comparison of Taint and Detection Thresholds Between Eulachon and Rainbow Trout. Report Prepared for Eurocan Pulp and Paper Co., Kitimat, B.C.

Bernstein, B.B. and J. Zalinski. 1983. An optimum sampling design and power test for environmental biologists. J. Envir. Mgmt. 16:35-43.

Bligh, E.G. and W.J. Dyer. 1959. A rapid method of total lipid extraction and purification. Can. J. Biochem. Physiol. 37: 911-917.

Bolger, T. and P.L. Connolly. 1989. The selection of suitable indices for the measurement and analysis of fish condition. J. Fish. Biol. 34:171-182.

Coad, B.W. 1995. Encyclopedia of Canadian Fishes. Canada Museum of Nature and Canadian Sportfishing Productions Inc. 928 pp.

Cone, R.S. 1989. The need to reconsider the use of condition indices in fisheries science. Trans Amer. Fish. Soc. 118:510-514.

CPAR. 1978. Protocol for fish tainting bioassay. Report prepared by Domtar Research Centre for Environment Canada, Environmental Protection Service, CPAR Project Report 775-1.

Draper, N.R. and H. Smith. 1981. Applied regression analysis. 2nd ed. John Wiley & Sons, New York.

Day, R.W. and G.P. Quinn. 1989.Comparisons of treatments after and analysis of variance in ecology. Ecol. Mong. 59:422-463.

Environment Canada. 1992a. Reference Method for the Determination of Polychlorinated Dibenzo-para-dioxins (PCDDs) and Polychlorinated Dibenzofurans (PCDFs) in pulp and Paper mill Effluents. Environment Protection Series Report EPS 1/RM/19, February 1992.

Environment Canada. 1992b. Internal Quality Assurance Requirements for the Analysis of Dioxins in Environmental Samples. Environment Protection Series Report EPS1/RM/23, October 1992.

Environment Canada. 1997a. Aquatic Environmental Effects Monitoring. Requirements at Pulp and Paper Mills and Off-Site Treatment Facilities.

Environment Canada 1997b. Pulp and Paper Aquatic Environmental Effects Monitoring Requirements. Annex 1 to EEM/1997/1.

Environment Canada 1997c. Fish Survey Expert Working Group report. EEM/1997/6.

Environment Canada and Department of Fisheries and Oceans. 1993. Technical Guidance Manual for Aquatic Environmental Effects Monitoring at Pulp and Paper Mills. EPS.1/RM.

FAU Update 94-1, Ontario Ministry of Natural Resources Fisheries Assessment Unit Newsletter. 1994. Nearshore community index netting (NSCIN): indexing the abundance of the warmwater fish community. Lake Simcoe Fisheries Assessment Unit, Sutton West, Ont. 2pp.

FAU Update 94-2, Ontario Ministry of Natural Resources Fisheries Assessment Unit Newsletter. 1994. Spring littoral index netting (SLIN): indexing the abundance of the coldwater fish community. Muskoka Lakes Fisheries Assessment Unit, Bracebridge, Ont. 2pp.

Folch, J., M. Lees and G.H. Sloane Stanley. 1957. A simple method for the isolation and purification of total lipids from animal tissues. J. Biol. Chem. 226: 497-509.

Fritz, E.S., W.H. Meredith, and V.A. Lotrich. 1975. Fall and winter movements and activity level of the mummichog Fundulus heteroclitus, in a tidal creek. Chesapeake Sci. 16: 211-215.

Gibbons, W.N., M.D. Paine, W.M. Gibson, G.A. Vigers and K. Kingston. 1992. 1991 operational monitoring survey of the Lesser Slave River. Volume III. Prepared for Slave Lake Pulp Corporation, Edmonton, Alberta by EVS Consultants, North Vancouver, B.C.

63 pp. + Tables + Figures + App.

Gibbons, W.N., M.D. Munn and M.D. Paine. 1993. Guidelines for monitoring benthos in freshwater environments. Prepared for Environment Canada, North Vancouver, B.C. by EVS Consultants, North Vancouver, B.C. 81 pp.

Gibbons, W.N., K.R. Munkittrick and W.D. Taylor. 1998a Suitability of small fish species for monitoring the effects of pulp mill effluent on fish. 1. Response of spoonhead sculpin exposed to bleached kraft mill effluent. Environ. Toxicol. Chem. In Press, NWRI Publication number 97-26a

Gibbons, W.N., K.R. Munkittrick and W.D. Taylor. 1998b. Suitability of small fish species for monitoring the effects of pulp mill effluent on fish. 2. Comparison of trout perch responses to white sucker exposed to bleached kraft mill effluent. Environ. Toxicol. Chem. In Press. NWRI Publicaiton number 97-26b

Goldstein, R. 1989. Power and sample size via MS/PC-DOS computers. Amer. Statist. 43:253-260.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley & Sons, Toronto. 257 pp.

Green, R.H. 1989. Power analysis and practical strategies for environmental monitoring. Envir. Research 50:195-205.

Grubbs, F.E. 1969. Procedures for detecting outlying observations in samples. Technometrics 11:1-22.

Harris, R.J. 1975. A primer for multivariate statistics. Academic Press, New York.

Hubert, J.J. 1980. Bioassay. Kendall/Hunt Publishing Co., Toronto, ON. 164 pp.

Hunn. J.B. 1988. Field assessment of the effect of contaminants on fishes. U.S. Department of the Interior, Fish and Wildlife Service Biological Report 88(19). 27 pp.

International Standards Organization (ISO). 1992. Sensory Analysis: Vocabulary. International Standards Organization, Geneva, Switzerland. ISO 5492.

Iredale, D.G. and R.K. York. 1976. Purging a muddy-earthy flavor taint from rainbow trout (*Salmo gairdneri*) by transferring to artificial and natural holding environments. J. Fish. Res. Board Can. 33: 160-166.

Jackson, D.A., H.H. Harvey, and K.M. Somers. 1990. Ratios in aquatic sciences: statistical shortcomings with mean depth and morphoedaphic index. Can. J. Fish. Aquat. Sci. 47:1788-1795.

Jenkins, R.E. and N.M. Burkhead 1993. Freshwater fishes of Virginia. American Fisheries Society, Bethesda, Maryland. 1079 pp.

Johnssen, P. and C.A. Kelly. 1990. A technique for the quantitative sensory evaluation of farm-raised catfish. Journal of Sensory Studies 4:189-199.

Klemm, D.J., P.A. Lewis, F. Fulk and J.M. Lazorchak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. EPA 600/4-90/030. U.S. Environmental Protection Agency, Environmental Monitoring Laboratory, Cincinnati, OH.

Kosmark, J.J. 1987. Standardizing sensory evaluation methods for marketing fish products. In: D.E. Kramer and J. Liston (Eds.). Seafood Quality Determination. Elsevier Science Publishers, Amsterdam. pp. 99-107.

Leblanc, J. et C. M. Couillard. 1995. Description de la période de reproduction d'un poisson sentinelle: le choquemort (Fundulus heteroclitus) de l'estuaire de la Miramichi Rapp. tech. can. sci. halieut. aquat. 2057: vii + 32p.

Leblanc, J, C.M. Coulliard, and J.F. Brethes. 1998. Modifications of the reproductive period in mummichog (Fundulus heteroclitus) living downstream from a bleached kraft pulp mill in the Miramichi Estuary, New Brunswick, Canada. CJFAS (in press)

Lindsay, R.C. and T.P. Heil. 1992. Flavor tainting of fish in the Upper Wisconsin River caused by alkyl- and thiophenols. Water Sci. Technol. 25: 35-40.

Lovell, T. And L.A. Sackey. 1973. Absorption by channel catfish of earthy-musty flavor compounds synthesized by cultures of blue-green algae. Trans. Am. Fish. Soc. 102: 774-777.

Malcolmson, L.J. 1995. Organization of a sensory evaluation program. In: Methods to Assess Quality and Stability of Soils and Fat-containing Foods. AOCS Press, Champaign.

Mann, H. 1951. Zur Frage der Geschmacksbeeinflussung durch Phenole. Fischwirt 1:164-165.

Martin, J.F. and I.H. Suffet. 1992. Chemical etiologies of off-flavor in channel catfish aquaculture. Wat. Sci. Tech. 25:73-79.

Meilgaard, M., G.V. Civille and B.T. Carr. 1991. Sensory Evaluation Techniques. CRC Press, Inc., Boca Raton, FL.

Miller, R.G. 1986. Beyond ANOVA: Basics of Applied Statistics. Wiley Publishers, New York.

Munkittrick, K.R., 1992. A review and evaluation of study design considerations for site-specificity in assessing the health of fish populations. J. Aquat. Eco. Health 1:283-292.

Muñoz, A.M., G.V. Civille and B.T. Carr. 1992. Sensory Evaluation in Quality Control. Van Nostrand Reinhold, New York.

National Council of the Paper Industry for Air and Stream Improvement, Inc. (NCASI). 1987. Guidelines to Methods for the Analysis of the Effects of Effluent on the Flavor of Fish. Technical Bulletin No. 513. National Council of the Paper Industry for Air and Stream Improvement Inc., New York.

Nelson, J.S. and M.J. Paetz. 1992. The fishes of Alberta. The University of Calgary Press, Calgary, Alberta.

Persson, P.-E. 1984. Uptake and release of environmentally occurring odorous compounds by fish. A review. Water Res. 18: 1263-1271.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. Can. J. Fish. Aquat. Sci. 47:2-15.

Peters, R.H. 1983. The ecological implication of body size. Cambridge University Press, New York. 329 pp.

Poste, L.M., D.A. Mackie, G. Butler and E. Larmond. 1991. Laboratory methods for sensory analysis of food. Agriculture Canada Publication 1864/E.

Randall, R.C., H.Lee II, R.J. Ozretich, J.L.Lake and R.J. Pruell. (1991). Evaluation of selected lipid methods for normalizing pollutant bioaccumulation. Environ. Toxicol. Chem. 10:1431-1436.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Fish. Res. Bd. Can. Bull. 191.

Roberts, W.E. 1988. The sculpins of Alberta. Alberta Naturalist 18:121-127.

Scott, W.B. 1967. Freshwater Fishes of Eastern Canada. University of Toronto Press, Toronto, Ontario, Canada. 137 pp.

Scott, W.B. and E.J. Crossman, 1973. Freshwater Fishes of Canada. Fish. Res. Board. Can. Bull. 184. Ottawa, Ontario, Canada. 966 pp.

Snedecor, G.W. and W. G. Cochran. 1971. Statistical Methods. Sixth Edition. Iowa State University Press.

Snedecor, G.W. and W.G. Cochran. 1980. Statistical Methods, 7th ed. Iowa State University Press, Ames, Iowa. 507 p.

Shuter, B.J. (1990). Population-level indicators of stress. Amer. Fish. Soc. Symposium 8:145-166.

Skalski, J.R. and D.H. McKenzie. 1982. A design for aquatic monitoring programs. J. Envir. Mgmt. 14:237-251.

Swanson, S.M. 1993. Wapiti/Smoky river ecosystem study. Prepared for Procter and Gamble Ltd./Weyerhaeuser Canada Ltd., Grand Prairie, AB by Sentar Consultants Ltd., Calgary, AB. 176 pp.

Taylor, L.R. 1961. Aggregation, variance and the mean. Nature (Lond.) 189:723-735.

Thaysen, A.C. 1936. The origin of an early or muddy taint in fish. I: The nature and isolation of the taint. Ann. Appl. Biol. 23:99-104.

Trippel, E.A. and H.H. Harvey. 1991. Comparison of methods used to estimate age and length of fishes at sexual maturity using populations of white sucker (*Catostomus commersoni*). Can. J. Fish. Aquat. Sci. 48:1448-1459.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell and C.E. Cushing. 1980. The river continuum concept. Can. J. Fish. Aquat. Sci. 37:130-137.

Vetemaa, M., L. Forlin and O. Sandstrom. 1997. Chemical industry effluent impacts on reproduciton and biochemistry in a North Sea population of viviparous blenny (Zoarces viviparous). J. Aquat. ecosys. Stress & Recovery 6:33-41.

Yantis, J.E. (Ed.) 1992. The Role of Sensory Analysis in Quality Control. ASTM Manual Series: MNL 14. American Society for Testing and Materials, Philadelphia, PA.

Zar, J.H. 1984. Biostatistical Analysis, 2^{nd} ed. Prentice-Hall, Englewood Cliffs, NJ.

6 Effects on Fish Habitat: Invertebrate Community Survey

The stated objectives of an Invertebrate Community Survey (ICS) for EEM (Environment Canada and Department of Fisheries and Oceans 1992, 1993, 1995 and Environment Canada 1997a) are to delineate the magnitude and spatial extent of habitat degradation due to organic enrichment or other forms of contamination, and to provide an evaluation of the aquatic food resources available for fish selected as a sentinel species in the fish survey (Chapter 5). However, without a direct comparison between fish diet and the benthic invertebrate fauna, the ICS is mainly aimed at examining habitat degradation. Therefore, the goal the benthic community survey is to determine structural differences (i.e. abundance, number of taxa, shifts in the kinds and dominance of taxa) in invertebrate communities in the vicinity of the mill effluent discharge. For freshwater, estuarine and marine mills, different design considerations will be required and these are addressed in Section 6.1. It is also recognized that benthic invertebrate surveys will not always use the same sampling design as the adult fish, sentinel species, or water quality surveys because of different criteria and challenges inherent in the different sampling protocols. It is critical though, for all mills, that sediment contamination surveys (Chapter 7) be done at the same times and locations as the ICS (see also section 6.2.1.5 on physical factors to be measured concurrent with the ICS). The objective of this chapter is to provide guidance on the study design and interpretation of results from an ICS in relation to EEM requirements. Specifically, this document elaborates on:

1) study design considerations (section 6.1),
2) standardization of methodologies (section 6.2),
3) data analyses appropriate to the study design (section 6.3).

The approaches recommended by this document are meant to be objective and utilize standard methods or techniques which will identify differences in communities between sampling areas. However, it should be recognized that not all site-specific considerations can be dealt with here and that consultation between mills and Regional Authorization Officers (RAOs) will be necessary to ensure site-specific considerations are taken into account. The recommended study design and reporting requirements are meant as minimum; additional sampling, data analysis and interpretation are encouraged. Furthermore, although most sections in this document refer to appropriate reference material, investigators are encouraged to review Green (1979; 1980), Resh and Rosenberg (1984), Holme and McIntyre (1984), Burd *et al.* (1990), Diaz and Schaffner (1990), Rees *et al.* (1990), Coull and Chandler (1992), Gray *et al.* (1992), Diaz (1992), a series of articles in Rosenberg and Resh (1993), Stanford and Spacie (1994), and Loeb and Spacie (1994) for information on the fundamentals of monitoring benthos as well as information which may be useful for interpreting survey data.

6.1 Study Design

The study design of an ICS is partially dependent on whether the mill in question is a new mill, considered at the Cycle 1 stage, or a Cycle 2 (or subsequent) mill. These two surveys differ in purpose and availability of historical information, but both are to be designed and conducted within the zone of effluent mixing and within an appropriate reference area (or areas) and, in most cases, will involve a minimum of one sample collection survey during each cycle. Considerations for the basic Cycle 1 design are discussed in Section 6.1.3. These requirements are considered the minimum as stipulated by Annex 1 but consideration should be given to going beyond the minimum since establishing a comprehensive baseline of information in regards to the benthic invertebrate community before operation will only be possible before effluent discharge commences. All Cycle 2 (or subsequent) mills, having conducted a Cycle 1 survey, will have some historical information to examine and are required to apply the Benthic Decision Tree (Figure 6.2) in consultation with the RAO to their Cycle 1 ICS results. This procedure for determining study designs for subsequent cycles is discussed fully in Section 6.1.4.

The Benthic Community Expert Working Group (BCEWG) report (Environment Canada 1997b) examines and reports on the results of the Cycle 1 invertebrate community surveys conducted at the participating mills and provides commentary and recommendations for EEM Cycle 2 (or subsequent) ICSs. The four basic study designs, developed by the BCEWG and outlined in Section 6.1.5.3, are directed at dealing with confounding factors identified. Further information in regards to dealing with confounding factors is in Table 6.2 of the decision tree and in the Reference Site Selection Section in Chapter 3.

6.1.1 Standard Nomenclature

Agreement on standard nomenclature regarding replicate sample site and area is a critical aspect with regards to QA/QC guidelines especially if the objectives of the overall program, are to carry out national or regional assessment of the environmental performance of mills. It is urged here, as it is in the Sampling Design Section that standardization of the terminology applied to sampling designs be consistent. This may also apply to other aspects where there are regional or cross discipline (i.e. fish vs. invertebrate) historical differences in nomenclature. Please refer to the Glossary in this document if clarification of some terminology is needed.

6.1.2 Sampling Locations

This section defines the terms **area**, **replicate station**[1] and **field sub-sample**[2]; terms which are used throughout the report and are schematically represented in Figure 6.1. Note that these definitions were derived with the assumption that the objective of an ICS is to detect differences between or among areas. The basic ANOVA study design (i.e., Control/Impact), outlined in detail in Section 6.1.5.3, includes reference and exposure areas, each consisting of a number of replicate stations, (i.e., the *replicates* for the ANOVA). Guidance for calculating the number of replicate stations required to provide a given level of power for an ANOVA design is provided in Section 6.1.5.5 and by Lowell (1997). However, in addition to determining the total number of sampling stations in each area, multiple field sub-samples are recommended within each station. The objective of multiple field sub-samples at each replicate station is to ensure that the sampling effort will produce an accurate reflection of the metrics of interest for the invertebrate community. Although the minimum number of field sub-samples required at each replicate station is currently 2, a detailed discussion of methods for determining the appropriate level of effort to apply at the field sub-sample scale is provide in Section 6.1.5.4. Because the level of replication for the ANOVA study designs is at the station level individual field sub-samples can either be processed in the laboratory separately or after pooling; but see Section 6.1.5.4 for a more detailed discussion regarding the use of information which may be gained from individual field sub-samples within a replicate station.

AREA: A sampling area is defined to be a relatively large geographical area which has a defined level of exposure to mill effluent. For rivers and streams, the extent of this area is defined in terms of the stream morphology and should encompass a length of river which includes at least three full meanders. This equates to a river length of 36X the bankfull width (Newbury 1984). For lakes and marine environments an area should be at least 100m x 100m, and large enough to adequately accommodate the necessary number of replicate stations. For the EEM Invertebrate Community Surveys, there may be three defined areas; one Reference area and two areas which have different exposures to pulp mill effluent; a near-field area and a far-field area Reference areas are to be outside the range of effluent exposure (i.e., a homogeneous, undetectable level of mill effluent) or have the least potential effluent exposure. Where possible, near-field areas will be $\geq$ 1% dilution of effluent and far-field areas will be $\leq$ 1% dilution. More guidance for the appropriate selection of Reference, Near-field and Far-field areas is provided in Chapter 2 and 3.

REPLICATE STATION: A replicate station is a specific, fixed sampling location within an area, and within which a sample consisting of a minimum of two field sub-samples will be collected. A replicate station is an entity which can be recognized, re-sampled and defined quantitatively (e.g., latitude and longitude). In lake and marine environments, each of the replicate stations will be a defined location with re-visitable dimensions (e.g. 10m x 10m). In rivers, a replicate station should encompass a longitudinal stretch of the river which includes one pool/riffle sequence, a river distance of 6X the bankfull width (Leopold *et al.* 1964, Newbury 1984, Leopold 1994). If this length of river is not feasible for a large river, an alternative definition would be similar to that suggested for lakes (i.e. a location with re-visitable dimensions (e.g. 10m x 10m)). Each replicate station will have a similar type of habitat within a given ICS.

---

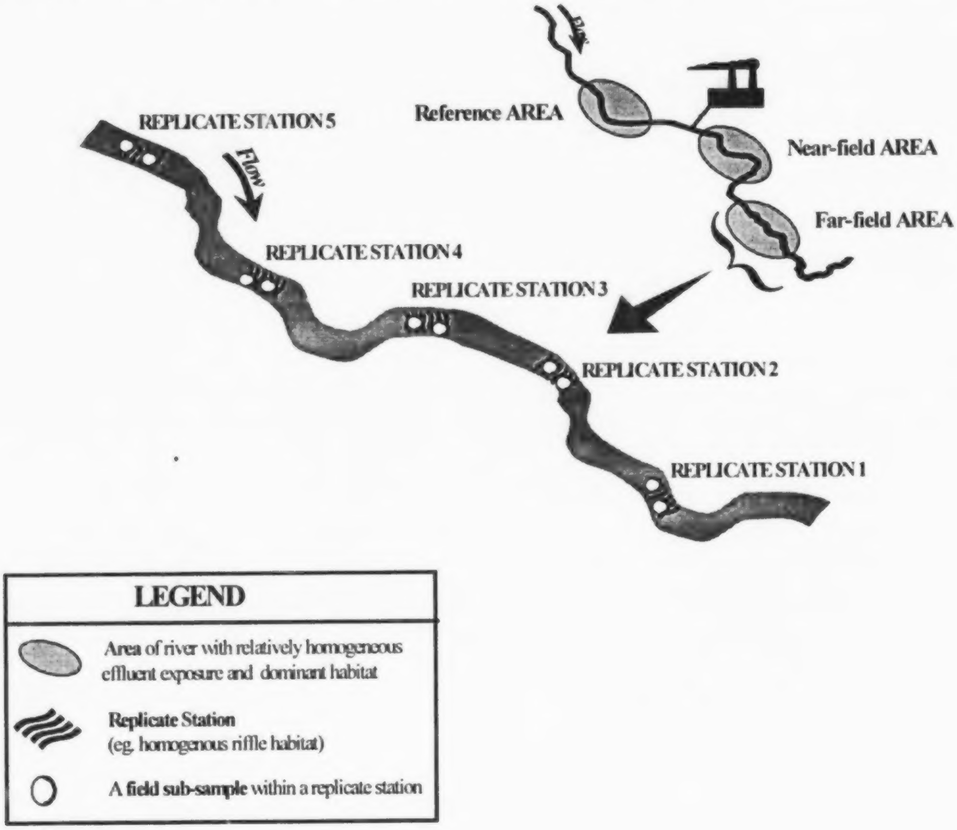
[1] Annex 1 (Environment Canada 1997a) refers to "replicate stations" as stations

[2] Annex 1 (Environment Canada 1997a) refers to "field sub-samples" as "sampling units"

FIELD SUB-SAMPLE : A field sub-sample will consist of an individual area or time-limited collection (e.g., one grab, core, cylinder, kick or U-net sample). Field sub-samples within a replicate station are to be collected in a random or stratified random pattern such that each point in the replicate station or habitat has an equal probability of being sampled (see discussion below). As noted above, each field sub-sample from a replicate station may be preserved and processed separately or pooled for processing. Regardless of whether the field sub-samples are pooled for processing; for statistical analysis, the simple ANOVA designs require the pooling of all the field sub-sample estimates from each replicate station into one number representing the value of the metric for the given replicate station.

The remaining consideration is in how to specifically allocate the field sub-samples within each replicate station, as there will likely be some spatial limit to the size of the available sampling replicate stations. As indicated above, placement of field sub-samples within replicate stations are to be considered "random". Random sampling requires that each member of the population of samples has an equal chance of being selected (Zar 1984). Sampling of benthic communities is often conducted in a "haphazard" fashion where there is no conscious bias (Klemm *et al.* 1990). This should not be confused with random sampling. One satisfactory method of random sampling is to number the universe or entire set of field sub-samples available from which the sample will be selected (Klemm *et al.* 1990). This can be accomplished by marking off equal distances along a transect and numbering each mark, or by dividing a section of a water body (the sampling replicate station) into grids and numbering each intercept. The total number of marks or intercepts is the total number of *potential* field sub-samples. Then, from a table of random numbers found in any statistics textbook, select as many random numbers as there are field sub-samples. Note, however, that as stated above, the habitat sampled should be the dominant habitat class and be consistent between areas and replicate stations. Therefore, when constructing a replicate station "grid" for determining the placement of random samples, the entire area of the grid should fall within the predefined acceptable habitat class. If specific locations within the replicate station grid are determined *a-priori* to be outside the range of acceptable habitat (e.g. a large log or boulder lies in the middle of a cobble riffle), then these sections of the grid can be excluded from the random allocation of samples. This ensures that the samples are placed randomly amongst the set of acceptable habitat locations available within the replicate

Figure 6.1 Examples of **AREA, REPLICATE STATION** and **FIELD SUB-SAMPLE** spatial scales for a basic Control/ Impact design (see Section 6.1.5.3)


Flow
Reference AREA
Near-field AREA
Far-field AREA
REPLICATE STATION 5
Flow
REPLICATE STATION 4
REPLICATE STATION 3
REPLICATE STATION 2
REPLICATE STATION 1
LEGEND
Area of river with relatively homogeneous effluent exposure and dominant habitat
Replicate Station (eg. homogenous riffle habitat)
A field sub-sample within a replicate station

station grid. Klemm *et al.* (1990) provide a good discussion of how random sampling can be conducted in a variety of habitats, while the reasons for random sampling are described in all introductory statistics textbooks.

6.1.3 Pre-operational and EEM Cycle 1 Invertebrate Community Surveys

New mills are required to complete the EEM pre-design and Cycle 1 Invertebrate Community Survey (ICS) prior to the start-up of operations. This first cycle of EEM is intended to characterize the benthic communities in major habitats which may be affected by mill effluent and to establish a baseline against which data from future cycles can be compared. This initial cycle will also allow for a critical assessment of the need to refinement the study design in future cycles or of the need for the introduction of alternative monitoring techniques. To address the stated objectives of the ICS for Cycle 1 mills, guidelines for minimal study design are presented below. Once a new mill begins production and discharges effluent, the mill will have four years to complete the following cycle and file a report with Environment Canada.

One specific objective of a Cycle 1 survey is to define areas that are relatively homogeneous in terms of habitat class and have specific ranges in level of exposure to mill effluent. Because, effluent is not yet being discharged, the delineation of the zone of effluent mixing will be estimated based on design criteria and the appropriate modeling techniques outlined in Section 2.3. This pre-design information will assist in the physical allocation of sampling areas and sampling stations within these areas. To fulfill the requirements of Annex 1 for Cycle 1 mills, a minimum of three sampling areas are to be included in the survey; Reference, Near-Field and Far-Field areas. The criteria for designation of these areas are based on the predicted levels of effluent in the receiving environment (in Chapter 2). This is the basic Control/ Impact Study Design discussed below (Section 6.1.5.4) but, depending on the site specific conditions, any one of the study designs may be applicable to an ICS for Cycle 1 mills.

During the study design phase of the ICS, the primary difference between Cycle 1 and 2 mills is the availability of invertebrate data gathered during a previous cycle. This information can be used in a power analysis exercise to determine more appropriate sample sizes for subsequent cycles. In the absence of preliminary information on site-specific variability at the reference area, the best generic sample sizes for a Cycle 1 survey were derived through a power analysis exercise which substituted effect size as a factor of SD. This information on power analysis provided in Chapter 3, indicates that a basic Cycle 1 C/I design, would include a minimum of 5 replicate stations/area, where the stations are considered the replicates and 2 samples at each replicate station which are, statistically, considered sub-samples. This sampling design provides a probability of Type 1 errors of 0.05, Type II errors of 0.20 (Chapter 3) and a critical effect size of 2SD (Section 6.1.5.5).

Thus, the first EEM cycle requires that the survey consist of the following:

1) at least three study areas, reference, near-field and far-field,
2) at least five replicate stations , in each of the three study areas; and,
3) a minimum of two field sub-samples to be taken at each station.

Note that, without *a priori* information on invertebrate density and variability within a station, the number of field sub-samples required to accurately reflect the true density at each station is arbitrarily set at two (but see section 6.1.5.5 for more information on allocation of sampling effort). This amounts to a total recommended sampling effort for Cycle 1 mills of 30 benthic samples. It is felt that this sampling effort and distribution would improve the initial information obtained by Cycle 1 mills. Where study designs other than the C/I design are appropriate, the same minimum sampling effort should be used, although the distribution of areas, stations and samples may differ.

A further recommendation for Cycle 1 surveys, is that the stations be located such that only the dominant habitat class (see Section 6.1.5.3) is sampled. This is a change from the original Cycle 1 requirements where all habitat classes were to be sampled in proportion to their abundance. Restricting sampling to the dominant habitat class allows the initial ICS to have more statistical power since the minimum sampling effort is not spread out over many different habitats. However, because this recommendation restricts the sampling effort and, the dominant habitat class will then be sampled consistently between areas and cycles in the future, it is paramount that the initial effort in determining habitat classes during the pre-design studies be diligent. Study areas that have extremely heterogeneous habitats or two habitats which are equally dominant may require a greater sampling effort than the minimums suggested above. Further increases in sampling effort, beyond the minimums required for Cycle 1, are recommended and could include any of the following: addition of one or more reference areas, addition of a far-far field area, addition of more stations per area, or the addition of more field sub-samples per station. Increases in sampling effort should be determined in consultation with the RAO.

6.1.4 EEM Cycle 2 Invertebrate Community Surveys

The intent of subsequent EEM cycles is to compare the results of the previous cycle, help refine monitoring techniques as needed, determine the factors leading to any detected effect, and document the spatial extent of any observed effect. Once Cycle 1 has been completed, its information is input into the Benthic Survey Decision Tree (Figure 6.2) which is a guide to the design of subsequent cycles for freshwater and marine environments. Note, that if historical information is available, it can similarly be input into the Benthic Decision Tree to aid in study designs.

Four basic study designs are outlined in Section 6.1.5.4 are aimed at estimating the magnitude and spatial extent of a potential pulp-mill effluent related impact. They attempt to take into consideration factors and possible constraints related to the availability and spatial distribution of suitable control ("reference") areas and the spatial extent and heterogeneity of potential impact areas. It should be emphasized that these guidelines, although considered the most applicable generic designs available, are *not* an exhaustive list of the possible means and ways of conducting an ICS. It is assumed that each study leader has sufficient knowledge to apply these recommendations in a sound scientific manner and is able to determine if unique conditions exist which would warrant modification of the study designs with the RAO.

In some cases, however, the confounding factors found in Cycle 1 mills may be unresolvable through sampling redesign. Currently, as is evident from Table 6.6 of the decision tree, there are no approved alternative methods available.

Figure 6.2 Benthic survey decision tree

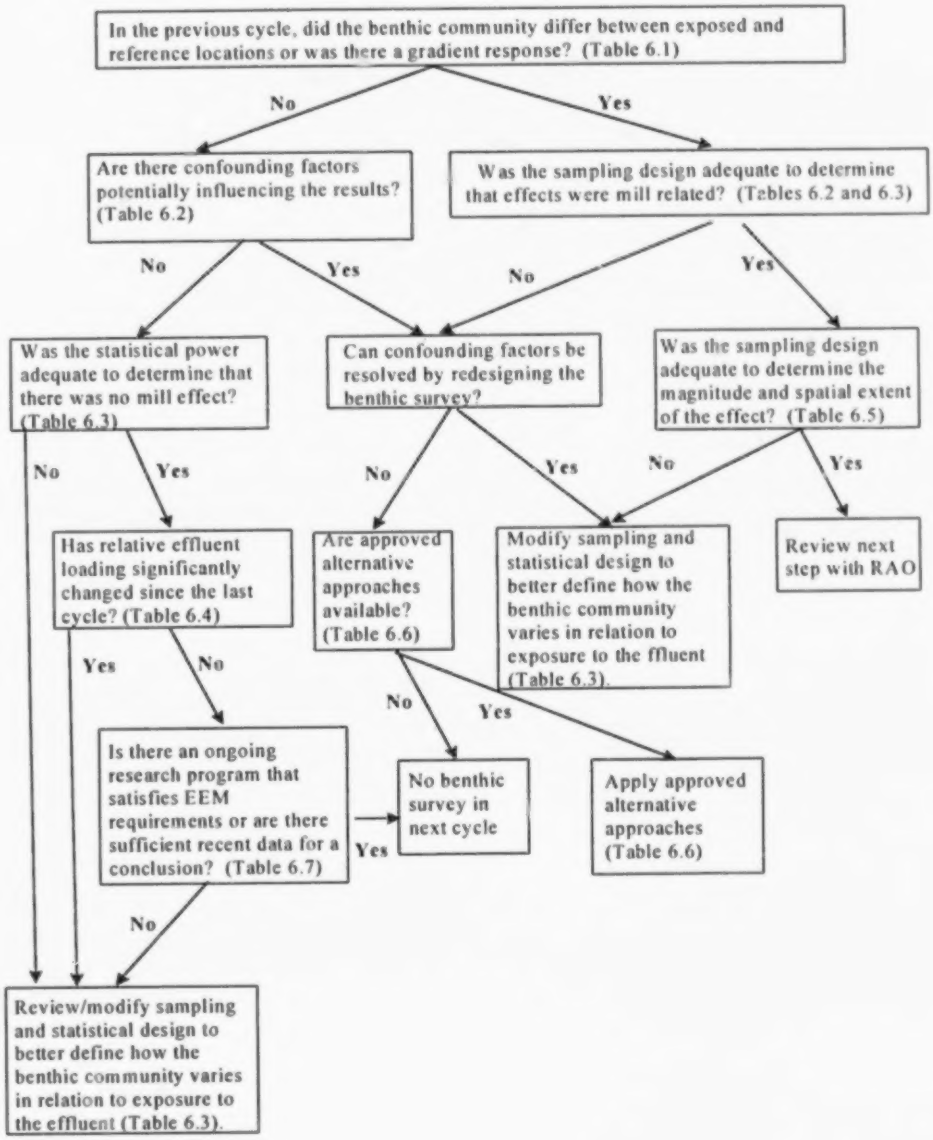

In the previous cycle, did the benthic community differ between exposed and reference locations or was there a gradient response? (Table 6.1)
No
Yes
Are there confounding factors potentially influencing the results? (Table 6.2)
Was the sampling design adequate to determine that effects were mill related? (Tables 6.2 and 6.3)
No
Yes
No
Yes
Was the statistical power adequate to determine that there was no mill effect? (Table 6.3)
Can confounding factors be resolved by redesigning the benthic survey?
Was the sampling design adequate to determine the magnitude and spatial extent of the effect? (Table 6.5)
No
Yes
No
Yes
No
Yes
Has relative effluent loading significantly changed since the last cycle? (Table 6.4)
Are approved alternative approaches available? (Table 6.6)
Modify sampling and statistical design to better define how the benthic community varies in relation to exposure to the ffluent (Table 6.3).
Review next step with RAO
Yes
No
No
Yes
Is there an ongoing research program that satisfies EEM requirements or are there sufficient recent data for a conclusion? (Table 6.7)
Yes
No benthic survey in next cycle
Apply approved alternative approaches (Table 6.6)
No
Review/modify sampling and statistical design to better define how the benthic community varies in relation to exposure to the effluent (Table 6.3).

Table 6.1 In the previous cycle, did the benthic community differ between exposed and reference areas or was there a gradient response?

Answering this question is equivalent to comparing various descriptors of the benthic community (e.g., total abundance or biomass, taxonomic richness) among the two or more levels of exposure to the effluent. A relationship between the benthic community and the effluent may be a simple difference in a community descriptor between one or more reference areas and an exposed area, or it may be a gradient of variability in a community descriptor among stations which vary in their level of effluent exposure.

A weight-of-evidence decision must be made based on both statistical and non-statistical comparisons of stations with varying exposure levels.

Statistical comparisons are based on community descriptors calculated from each observation made at a station with a given level of exposure. An example is the average number of genera per replicate. Variation among stations with different exposure levels can be assessed with 1-way ANOVA or its non-parametric analogs. An ecologically significant effect size must be defined for each community descriptor subjected to this statistical test (Table 6.3).

Non-statistical comparisons are based on community descriptors calculated from the pooled observations made at a station with a given level of exposure. An example is the total number of genera found at a station. Variation among stations with different exposure levels can be assessed by inspection. If more than three levels of exposure have been sampled, the relationship between a community descriptor and exposure level can be statistically evaluated, albeit usually with weak power. Although statistical comparisons are usually not appropriate for these comparisons, an ecologically significant effect size must be defined for each community descriptor subjected to this comparison (Table 6.3).

A typical assessment, where several to many community descriptors are used, will result in a variety of judgments about the relationship between the benthic community and exposure to the effluent. For example, average number of genera per replicate may not be related to effluent exposure, but average total abundance per replicate is. In the case of a mixed result, the previously established response of community descriptors to mill effluent will be taken into consideration when constructing a weight-of-evidence decision. If descriptors known to be sensitive to the effluent are not correlated with exposure at a given mill, while other variables not known to be affected by mill effluent are correlated, then linking observed effects to the mill is less likely.

Table 6.2 Was the sampling design adequate to suggest that the effects were linked to mill effluent? Are confounding factors or sampling methods potentially influencing the results?

The interpretation of benthic community response may be influenced by confounding factors or application of incorrect sampling methods. There are four categories of such factors:

1. Environmental variables
2. Multiple discharges or historical effects
3. Time of sampling
4. Sampling methods

1. Environmental variables

It is possible that environmental variables such as depth and particle size can confound the interpretation of effects. For example, if depth increases in proximity to the mill, then it may not be possible to separate the effect of the mill from depth effects. If feasible, a change in sampling design by modifying stations at macro (e.g., geographic location), meso (e.g., riffle within a river area), and micro (e.g., substrata within a riffle) spatial scales may reduce the problem. Augmenting the design to better characterize reference conditions by sampling more reference areas (e.g., moving from a gradient to a radial gradient design, or a control-impact to a multiple-control-impact design; Table 6.8) might reduce the problem.

2. Multiple discharges or historical effects

A potential for confounding effects exists, if areas with varying levels of exposure to the mill effluent also have varying levels of exposure to other effluents or stresses, or habitat modifications from historical effects of the mill. If feasible, a change in sampling design by modifying sampling locations may reduce the problem. Core sampling may also be useful in depositional environments to resolve historical trends.

3. Time of sampling

The time of year or the particular year of sampling may confound the interpretation of community response to effluent. This can be assessed by knowledge of the phenology of benthic community species (i.e. relation between climate and life history characteristics) and examination of data collected in previous years from the reference areas.

4. Sampling methods

If standard methods have not been used, this may obscure any effect of the effluent on the benthic community.

Table 6.3 Was the statistical power adequate to determine that there was no mill effect?

If there is no relationship established between a community descriptor and the effluent, the power of the statistical tests should be assessed.

The **power** of a statistical test is the probability of detecting a given magnitude of relationship between the benthic community and the effluent, known as the effect size, when areas of varying levels of exposure to mill effluent are compared.

Effect is the magnitude of the relationship between the benthic community and the effluent that is considered ecologically significant.

Effect size of a particular community descriptor can be defined in three ways (Lowell 1997):

a) some function of ***variability*** in the community descriptor ***between observations at various exposure levels***
b) ***existing information (scientific literature)*** about the acceptable magnitude and nature of response of a community descriptor to the effluent
c) ***societal input*** into what constitutes an acceptable magnitude and direction of ecosystem change

Once ***effect size*** is defined, the adequacy of the sampling design can be assessed. The power of detecting the effect size then must be calculated. If the power was less than 80%, even though the results show no significant difference, then the study cannot conclusively show there was no relationship between community and effluent.

Table 6.4 Has relative effluent loading significantly changed since the last cycle?

The absence of a relationship between community structure and effluent is based on the relationship between the loading to and the conditions of the receiving system. If either of these changes **significantly** then the assumption of no impact is no longer valid, and a sampling program should be initiated in the following cycle .

The following changes in effluent or the receiving system could trigger such a sampling program depending on site specific conditions:

Relative effluent quality and quantity changes:

Ecosystem response is determined by the relative loading of contaminants (e.g., nutrients, toxics) to the receiving environment (eg. dilution capacity). Contaminant loading is determined by both concentration and volume of effluent in relation to flow or volume of the receiving environment. Therefore, if either of these change and the RAO considers the changes to be significant, then a sampling program is required.

Table 6.5 Was the sampling design adequate to determine the magnitude and spatial extent of the effect?

If a relationship between the benthic community and the effluent has been established, but the spatial extent of the relationship is not definable from the data collected, increased sampling effort at intermediate levels of exposure is required.

Table 6.6 Are approved alternatives available for the invertebrate survey?

No specific alternatives are recommended at this time. However, various alternatives are under review and this Table will be updated as alternatives are considered.

Table 6.7 Are there sufficient recent data for a conclusion?

The purpose of the EEM program is to evaluate the adequacy of the Regulations. There are three possible conclusions which can be reached: there is no effect, there is a mill-related effect, or there are confounding factors which prevent determination of the cause of an effect.

1. For determining that there are or are not changes, a minimum of two cycles, with a power and effect size that satisfies the requirements of Table 6.3, are required.
2. Confounding factors - If there are sufficient data to demonstrate that other discharges or contaminant sources are primarily responsible for observed changes or absence of observed changes, and the survey cannot be re-designed to isolate mill effects, then there can be a conclusion that confounding factors prevent interpretation for the purposes of assessing the adequacy of the Regulations.

6.1.5 Sampling Design

This section deals with the specifics of sampling design for EEM invertebrate community surveys. Specifically, this section reviews the timing of sampling, recommended sampling designs and their applications and possible alternative approaches which are currently under investigation.

6.1.5.1 Selection of Reference Areas

Marine and Estuarine

1) The reference area should be within the same water body and hydrographic current or tidal regime as the exposure area. In other words, the closer the reference area is to the exposure area, the better. Benthic invertebrate communities are considerably higher in species richness, have more complex trophic relationships, faunal size ranges and reproductive strategies than in freshwater. This is simply because there is a much higher diversity of types of organisms in the marine benthos than in freshwater. Because of this complexity, and the multitude of interactions between species in marine benthos, small shifts in physical or chemical conditions can dramatically alter the overall benthic faunal community. Add to this the effect of increasing variation in chance larval settlement with increasing geographic distance (geographic "drift" in community structure) and physical barriers in complex coastlines, and it is very rare to find similar invertebrate communities from one bay or fjord to the next, or to be able to predict specific benthic community structure based on sediment factors (for recent review on invertebrate sediment interactions in the marine see Snelgrove and Butman 1994). In order to have some confidence that the "natural" benthic community is similar enough from one coastal area to the next, there must be sufficient water exchange between them. This is more likely in open coastal areas than in isolated bays and fjords.

2) Reference areas which are not in the same water body or hydrographic regime may only be suitable for comparisons of summary characters such as shifts in abundance or species richness. If the habitat conditions are similar enough to the exposure area, it may also be possible to compare larger-scale biotic factors, such as the presence of characteristic, long-lived depth/substrate specific taxa described by Thorson (1957) as "parallel communities". It is not feasible to identify reference sites external to the water body for the mill which are suitable community-level comparisons in marine or estuarine systems.

   The use of a gradient approach is recommended for any "static" (limited movement) communities sampled in the estuarine/marine, such that the reference data is the "least affected" by mill effluent, but still mainly within the range of physical conditions and circulation patterns affecting the mill environs.

3) The reference area must have very similar substrate type (particle size, sorting, natural chemistry), depth properties, current regimes, physical water properties, nutrient regimes and drainage characteristics as the mill survey area. Some special considerations are important for determining the suitability of reference areas for marine and estuarine mills. Physical factors in the estuarine/marine environment which tend to be more complex than in freshwater include; salinity, tides and currents, and sediment sulfides. For intertidal areas slope, wave exposure, light and tidal exposures are also important. In order to determine if a reference area is physically suitable or not, referral to detailed classification systems as outlined for marine shorelines to deep coastal areas include Frith *et al*. 1993, Booth *et al*. 1996, Robinson and Levings (1995), Hay *et al*. (1996) and Robinson *et al*. (1996). Specifically, estuarine classification has been reviewed by Matthews (1993), Scott and Jones (1995), Finlayson and van der Valk (1995) and Levings and Thom (1994). In the US, the most widely used system is that of Cowardin *et al*. (1979), Cowardin and Golet (1995)with expansions proposed by other authors.

4) The reference area must not be confounded by the presence of strong physical characters or other discharge factors not present at the mill survey area. These can include: historical effluents which may still be affecting benthic habitats; present-day effluents such as other mills, sewage or stormwater discharge, industrial discharge, agricultural run-off, etc; physical factors such as dams or log-booms; extraordinary siltation or clogging from logging, periodic flooding, etc. Cycle 1 results showed that it is difficult to find suitable reference areas within the same water body or drainage basin in marine and estuarine areas which can be assured of being outside the influence of the mill or any other anthropogenic effects.

Freshwater Rivers and Lakes

1. The following information should be considered when defining and delineating reference areas in freshwater. **Using maps, locate area of interest** (i.e. mill site) and identify the surrounding
   - Ecoregion
   - Drainage Basin with the Ecoregion
   - Land-use & Vegetated Buffer Strip Adjacent to Water.

If a drainage basin occupies more than one ecoregion, reference areas are to be selected so that there are a proportional number in each ecoregion (i.e., a MCI design may be necessary). The size of the drainage basin selected is based on stream order. If a mill site is located on a fourth order stream, the drainage basin should be delineated at a fifth order stream (i.e. at the junction of two fourth order streams).

2. Assessment of Study Area:

- the characteristics of the effluent field should be understood (i.e., initial dilution, dispersion etc.)
- identify the number and location of exposure areas (near field, far field, far-far field)
- determine if there are other significant inputs in the study area.
- identify the receiving water habitat

3. Type of Receiving water - Rivers:

1 Consult section 6.1.5.3, Sampling Locations and Section 6.1.5.5., Sampling Effort, for definitions of areas and replicate stations and number of stations per area

2. a) If upstream inputs are absent, select reference area(s) within the drainage basin in which the mill is located.
b) If non-point or other point source inputs occur upstream of the effluent, select reference area(s) in the nearest comparable drainage basin with minimal development that occurs within the same ecoregion.
c) If physical disturbance of the river valley is associated with the mill, effluent effects may be confounded by the disturbance. Accordingly, additional physically matched reference areas should be selected.

3 Conduct field inspection to characterize the exposure areas (s) with respect to:

- Valley Segment
- Stream Reach
- Channel Unit

4 Conduct field inspection so that reference area(s) are paired with characteristics of exposure areas with respect to:

- Valley Segment
- Stream Reach
- Channel Unit

It is possible that different exposure areas may occur in different valley segments. If so, reference areas should be selected with paired valley segments of each exposure site.

5 Final check (Hughes 1995). Reference areas should have:

- a natural channel
- stable banks or shorelines
- an extensive vegetative buffer strip
- a natural hydrograph

4 Type of Receiving Habitat - Lakes

1. Consult section 6.1.5.3, Sampling Locations and Section 6.1.5.5, Sampling Effort, for definitions of areas and replicate stations and number of stations per area

2. a) If non-point or other point sources are absent, determine sphere of influence originating from effluent. This is particularly important for lakes whose discharge or effluent flow is not unidirectional.
   b) If non-point or other point source inputs occur elsewhere on lake, select reference area(s) in the nearest comparable lakes within the same or adjacent drainage basin within the same ecoregion.
   c) If the mill effluent is associated with physical disturbance in the area, effluent effects may be confounded by the disturbance. Accordingly, additional physically matched reference areas should be selected.

3 a) If tracer studies indicate that effluent effects are local and restricted, select reference areas within the lake in which the mill occurs. These reference areas should occur in separate, but comparable bays or basins of the lake.
   b) If tracer studies indicate that the identified effluent is dispersed throughout the lake, select reference area(s) in the nearest comparable lake within the same or adjacent drainage basin within the same ecoregion.

4 Use maps and existing literature to characterize the lake:
- Origin, Location, Hydrological Linkages
- Morphometry
- Trophic State
- Zone, Substrate Type

5 Reference areas should match or pair characteristics of exposure sites with respect to:
- Origin, Location, Hydrological Linkages
- Morphometry
- Trophic State
- Zone, Substrate Type

6 Final check (Hughes 1995). Reference areas should have:
- natural shoreline structure
- stable shoreline
- extensive natural riparian vegetation
- natural water level.

6.1.5.2 Sampling Periods

The seasonal timing of sampling is to be consistent from survey to survey. The preferred period for sampling is when biological diversity is highest. This will allow the amount of biological information collected to be maximized and to increase the potential of observing the invertebrate communities during a period when response patterns are expected to contain the greatest biological information. In general, this will correspond with the seasonal recruitment cycles of benthic organisms. Reproductive periods and patterns of abundance of benthic species are related to climate and the abundance of food supply. Many insects with freshwater stages reproduce in the spring and fall. For most lotic habitat types, sampling is best conducted during the fall when the majority of taxa will be present and/or large enough to be collected by the sampling equipment. In large lakes where the benthic community is often dominated by annelids, crustaceans, and molluscs, insect emergence periods are of less importance in the timing of sampling. In marine environments, sampling should be conducted in late summer or fall as many benthic forms have planktonic larval stages that do not appear in the benthos until this time period and populations with spring recruits have stabilized by this time. If historical benthic surveys exist for the system under investigation, it would be useful to examine the data and, if appropriate, conduct the survey during similar periods so that the surveys can be compared. However, the preferred season for the survey may be limited by the sampling gear selected, flow conditions, and effluent conditions. All sampling gear may be difficult to use during periods of high flow and sampling during periods when effluent is not being discharged should be avoided. An understanding of the seasonal patterns and life cycles of the taxa along with changes in the hydrologic regime found within the specific system is needed to properly determine the appropriate timing for the survey. Rosenberg and Resh (1993), Johnson *et al.* (1993), Rees (1984), Malley and Reynolds (1979), Barber and Kevern (1974) and Jonasson (1955) provide information which may assist in the selection of a sampling period.

6.1.5.3 Selection of Dominant Habitat Type

A change from the previous technical guidance is that only the dominant habitat type should be sampled during an ICS (REF). The sampling of all habitat classes in proportion to their abundance (the previous recommendation), is probably not necessary for adequate biomonitoring. Similarly the US EPA (Plafkin *et al.* 1989) found in a pilot study, in rivers, that the macroinvertebrate communities in riffle habitats responded to impacts in a similar fashion as the multihabitat assemblages. For example, under stress, taxa richness was reduced by the same proportion in both the riffle and multihabitat assemblages. With limited resources available for most biomonitoring studies, the sampling of a single habitat is intended to reduced the variability which is inherent in sampling natural substrates but which would be even greater if the same sampling effort is spread over a range of habitat types. Therefore, for purposes of cost effectiveness and ease of data interpretation, a decision needs to be made to sample either depositional or erosional habitats in the freshwater receiving environment. In addition, decisions about sampling intertidal versus subtidal substrates in the estuarine/marine mills will depend on which is the appropriate receiving environment. This section provides some guidance on the approach and questions to consider in the selection of which habitat class should be sampled during an ICS.

Freshwater Receiving Environments

A critical aspect of the pre-design stage of an ICS involves identifying and locating appropriate habitat types. Without an accurate and current habitat inventory decision making in regards to dominant habitat type becomes difficult. For guidance in habitat inventory and classification refer to Chapter 2, Section 2.3. For additional guidance for stream environments refer to Cuffney *et al.* (1993), Plafkin *et al.* (1989) and Meador *et al.* (1993).

The selection of the appropriate habitat type to sample for an ICS, in freshwater receiving environments requires consideration of the following questions:

1) Which habitat, in the absence of human influences, supports the richest assemblages of invertebrates within the study basin?
2) In which habitat are the invertebrates most likely to be exposed to sediment or water borne contaminants for extended periods of time? and ,
3) Which habitat type is present in the highest proportion in the study basin?

If the answer to all three questions is the same habitat type the decision is a simple one. However, this is rarely the case. For example, in streams, the species rich riffle habitat is expected to be highly sensitive to water-quality changes because it can support a diverse assemblage of species that display a wide range of sensitivities to water-quality changes. Therefore, the community in this habitat has the potential for greater change than less species rich communities. However, these species rich riffles occur in erosional areas and may not be exposed to sediment borne contaminants in either high concentrations or duration. In contrast, the fauna of depositional areas, which are generally less rich taxonomically, are of interest during biomonitoring exercises because they are exposed to concentrations of sediment borne contaminants for longer periods. Consequently, communities in depositional areas may respond to contaminants before the more sensitive but less exposed riffle habitats. The third consideration, of course, is the extent of a particular habitat in the study basin. If a highly sensitive and diverse habitat only comprises 5% of the study basin, changes in the invertebrate community in this 5% may not be considered ecologically or socially relevant. It is beyond the scope of this document to provide advise on societal values. The fourth important question to consider is :

4) For which habitat class is there historical information available, either from a previous Cycle or other sampling programs?

Thus, during the design phase of an ICS, the selection of which habitat type to sample is critical and will affect many of the other study design considerations discussed throughout this document. Careful consideration must be given to questions raised in this section and their relation to the ecological effects which are expected . If more than one habitat type was sampled during the first cycle, the data should be examined to guide the design of the next cycle (see Chapter 3 for further discussion on critical effect size and power analyses). Note, that if it is deemed advisable given the above considerations that more than one habitat type be sampled this should be discussed with the local RAO. However, the addition of a second habitat type should not be at the expense of a reduced sampling effort in the first habitat type, at either the field sub-sample, station or area scale.

Marine/Estuarine Receiving Environments

The habitat which is most common spatially and most likely to be affected by the effluent is the dominant habitat. However, selection of major habitats is partially related to viability of sampling in the marine. For example, if the major habitat is a vertical rock cliff to a depth of 300 m immediately outside the effluent discharge, this is rarely feasible for benthic invertebrate sampling without extraordinary equipment. When multiple habitats are available and appropriate, some choices must be made. In some cases, more than one habitat may have to be sampled (radial gradient or similar design). Where there is a choice, sampling of soft substrates is preferred because methods are generally more quantitative. Regardless of the selection of habitat, it's use must be justified for and approved by the local RAO.

A list of questions or considerations for selecting the habitat to be sampled is given below.

1. What is the habitat which is most biologically active or "important"

When the subtidal environment most exposed to mill effluent consists of both solid and unconsolidated sediments, then either both substrates need to be sampled or a decision on which to sample must be taken. All other factors being equal, the unconsolidated sediment is the most efficient to quantitatively sample. However, when it is obvious that a course sand substrate is almost devoid of macrofauna within the top 10 cm (or the depth of penetration of a sampling device), whereas the nearby rocky reef is extremely rich and an obvious haven for many fish, it is the most "active". Similarly, if there is an important fisheries resource in one major habitat type which is directly exposed to mill effluent, it may be considered the most biologically important. The pre-survey information for most mills should include some knowledge on these considerations, and they must then be justified to the local RAO.

2. Can the dominant habitat be "classified" according to recognized physical type and characteristic species?

Habitat classification systems have been discussed and reviewed by many researchers and can be useful for comparing "expected" biotic factors with actual biotic factors present in the mill vicinity. Some relevent references for marine classification for worldwide shorelines to deep coastal areas include Frith *et al.* 1993a, Booth *et al.* 1996, Robinson and Levings (1995), Hay *et al.* (1996) and Robinson *et al.* (1996). Specifically, estuarine classification has been reviewed by Matthews (1993), Scott and Jones (1995), Finlayson and van der Valk (1995) and Levings and Thom (1994). In the US, the most widely used system is that of Cowardin *et al.*(1979), Cowardin and Golet (1995)with expansions proposed by other authors.

3. Is the effluent discharge depth and/or buoyancy most likely to affect the intertidal or subtidal regions?

If the effluent discharge is and remains mainly intertidal then this should be the targeted habitat. However, if the effluent affects both intertidal and subtidal habitats then the subtidal is the preferred habitat, because this most likely impacts fish. If suitable, both habitats may be sampled in consultation with the RAO.

4. What is the dominant shoreline habitat in the intertidal area which is DIRECTLY exposed to effluent discharge?

In many cases, coastal shorelines will be mixed silt, sand, gravel and rock substrates. In bays near freshwater discharge, there tends to be accumulations of soft substrate such as sand or silt. Estuarine mills will usually have dominantly soft substrates from river-borne material. If there are similar percentages of both depositional and erosional habitats, the preferred habitat to sample is depositional, because this type will accumulate the discharged material from mill effluents and is more likely to suffer deleterious effects from organic enrichment. Erosional substrates tend to be kept "clean" by high current action or scouring.

However, if the percentage of solid substrate habitat is much greater than soft substrates, or if a previously "clean", rocky shoreline has begun to accumulate sediment related to mill discharge, then this may be the preferred habitat to sample.

5. Which areas in the dominant habitat are exposed to confounding factors which may affect benthic communities?

Benthic communities in naturally or anthropogenically confounded sampling areas are problematic to use for interpreting effects from mills. Obviously they should be avoided. For example, in situations where solid and unconsolidated substrates are present, only one of these may be outside the influence of the confounding factors. That habitat should be chosen for sampling.

6. What is the environment affected by subtidal discharges?

Obviously, the environment most exposed to effluent should be the targeted sampling region and will also determine the type of sampling design used. In an estuary, if the discharge occurs at the surface where there is a strong and permanent surface freshwater layer with little intrusion of saltwater at high tide at depth, then the dominant habitat to be sampled is downstream from the mill. However, where there is a strong tidal intrusion sampling will have to go upstream and downstream. There are numerous other factors of this type to consider, all of which require detailed pre-survey site-specific information about the habitats and the pattern of effluent dispersal.

6.1.5.4 Sampling Program Designs

Four benthic sampling program designs are recommended for marine and freshwater community assessments:

1) Control/Impact design (C/I)
2) Multiple Control/Impact design (MC/I)
3) Simple Gradient design (SG)
4) Radial or Multiple Gradient design (RMG)

These designs fall into two basic categories with different philosophical approaches. These are; a) the Control/Impact or Multiple Control/Impact designs which are ANOVA type designs used to determine the magnitude of difference between homogeneous exposed and unexposed areas, and b) the Gradient (simple, radial or multiple) designs, which are intended to examine changes in community structure along a physical and/or effluent gradient.

Control/Impact design

The simplest study design for use in environmental effects monitoring is the Control/Impact (or reference-exposure) design (Green 1979). In rivers, this consists of one upstream reference area and a series of downstream exposure areas minimally representing near-field and far-field areas and, if required, (i.e., as determined in the benthic decision tree) a far-far-field area (Figure 6.3). Levels of exposure to mill effluent differ between areas but are homogeneous within each of these "areas". Habitat classes sampled are to be consistent between areas and, with the exception of exposure level, these areas are to be as similar as possible in as many traits as possible (for e.g. substrate, current velocity, water properties, environmental gradients, land use, etc.). This first study design is recommended for simple, homogeneous rivers and streams without confounding upstream or near-site discharges from other sources, or where the dominant habitat occurs in spatially discrete but homogeneous patches (i.e., intermittent rocky outcroppings)

Multiple Control/Impact design

Two of the major problems associated with the use of a single reference area are that it can be easily confounded by other factors and it is pseudo-replicated (Hurlbert 1984). For example, a single reference area may differ from an exposure area with respect to several environmental variables in addition to the point-source effluent. In systems where an appropriate reference area is not available due to confounding factors, or it is determined, after review of historical information (e.g., Cycle 1 data), that more reference areas are desirable, the multiple control/impact design may be used. Schematic diagrams of this design for application in either large rivers or coastal or lake-situated mills are presented in Figure 6.4. Sampling schemes should be devised so that reference sites with minimal degradation are located in adjacent watersheds or bays and that comparable habitats classes spanning the range of habitats found within the mill study area are selected.

The design philosophy of both the C/I and MC/I designs is that a specific difference in magnitude of effect between a series of areas is being examined. This lends itself to a classic ANOVA design with associated power analyses (Chapter 3). These methods are statistically tractable and may provide weight of evidence indicators as to whether or not there is a biological effect from the mill and the magnitude of that effect. These designs assumes that effluent exposure and habitat conditions are homogeneous amongst all stations within a sampling area or that effluent exposure within an acceptable range for a particular defined area (e.g., far-field ≤ 1%). Depending on the flow regime and habitat conditions the assumption of homogeneity mentioned above may not exist and a gradient design would be more appropriate.

Simple Gradient design

When the habitat is not homogeneous, such that environmental changes are evident along the effluent dilution path, gradient designs are recommended. Gradient designs are also most suitable for situations in which there are no suitable "upstream" reference areas available, or where a suitable separate reference area cannot be found in adjacent watersheds or bays (see Chapter 3 for guidance on selection of suitable reference locations).

Gradient designs are recommended for marine areas where there is a point source of effluents which declines with distance away from the source concurrently with changes in habitat parameters. Philosophically, the gradient approach examines variations from expected dynamic "patterns" of correlated biotic and environmental factors over spatial gradients. This is more suited to a regression type of analysis (or equivalent) in which replication (i.e., 5 stations within an area) is less appropriate than expending a similar effort to obtain accurate measurements of biotic and habitat variables over a sufficiently broad range of the gradient conditions. In this approach a point source discharge is expected to have a "declining" effect away from the source. Therefore, it is not always feasible to make the judgment that there either "is" or "is not" an effect at a given location, rather there is a declining gradient of effects away from the point source. At a certain point along the gradient it is necessary to judge that this effect is no longer measurable or important. Therefore, in simple gradient designs, reference information is obtained from the stations furthest away from the effluent source.

A gradient does not imply straight lines, or the even spacing of stations within areas. In fact, the lack of homogeneous sampling areas is a direct cause for choosing a gradient design. The spacing of samples may be more or less continuous on a gradient away from the discharge, with less emphasis on distinctly different dilution zones and more on adequate spatial coverage. There are no "blank" spaces between distinct sample areas, but rather a continuum of sampling stations along the gradient. However, if a change in effluent dilution within the receiving environment is abrupt, more sampling effort may be desirable over these stretches to accurately track rapid changes in mill effects (potentially both increasing or decreasing). The design should always attempt to answer the questions: What is the spatial extent of that effect? What is the magnitude of effect, if any?

The gradient approach is recommended for physically "linear" habitats such as some rivers, estuaries, intertidal shorelines and narrow coastal fjords, where there are no upstream or other suitable reference areas, and where a single gradient of stations provides sufficient coverage to determine spatial extent of effects. As with the other designs, minimally two randomly selected samples at each station are recommended (but see sampling effort - 6.1.5.4 below). However, it is important, as much as possible, to minimize the number of co-occurring gradients. Too many changing gradients along the sampling path will make interpretation of the "effecting" factors difficult. If unavoidable, such complex patterns will require detailed sampling of other factors which may be contributing to faunal changes along the gradient. For example, in intertidal surveys, the best design may be along a gradient away from the mill effluent source, but this will require adequately measuring changes in factors such as; substrate type, slope, light and wave exposure (see Chapter 7)

Radial or multiple gradient design

A fourth option, the radial gradient approach, is available for coastal and lake-situated mills. Sampling is conducted away from the effluent source in several gradient transects (Figure 6.6). As in the multiple C/I approach, the use of a radial gradient design will provide a broader range in reference conditions. Further, a better spatial coverage will be sampled which can be important in non-homogeneous, open coastal and lake shore areas, which often have complex current and circulation patterns or a variety of equally important major habitat classes or gradients. In some cases, it may also be useful to compare "affected" versus "reference" gradients. The most appropriate design for each EEM invertebrate community survey should be determined based on previous experience within the area (i.e. Cycle 1 information, historical data and/or pre-design information on habitat classes).

A summary of the attributes, applicability and limitations of the four sampling designs is presented in Table 6.8. It is emphasized that these four designs are recommendations, and not an exhaustive list of possible study designs. There may be combinations of these approaches which may be appropriate to answer the EEM requirements. One example of a combined strategy is given as follows: if the dominant habitat is spatially compact and homogeneous with distinct patches and an effluent gradient exists, but a suitable reference area cannot be located upstream or in adjacent watersheds, a modification of one of the designs may be appropriate. In this case, the C/I design could be modified, so that the reference area is "downstream" instead of "upstream" (Fig. 6.7) of the point source. The downstream reference area would of course have to meet the same reference area criteria as other designs. Another example is where the effluent becomes diffuse quickly (i.e., well mixed coastal or lake shore areas), but there is a pronounced depth gradient, a depth-stratified random sampling pattern may be used as described in Gray *et al.* (1992). However, if departure from these four basic recommendations is desired, consultation with the local RAO is necessary.

Table 6.8 Benthic sampling program designs

Design Type	Receiving Environment	Reference / "control" Areas	"Impact" Areas	Statistics
Control/Impact (C/I)	freshwater rivers and streams	a single reference area, upstream of mill discharge	Near field/Far-field/Far-far field (optional) as defined by effluent concentration in receiving environment	ANOVA
Multiple - Control/Impact (MCI)	freshwater, rivers or lakes with spatially homogenous lake shores	Multiple reference areas in the same or environmentally similar adjacent watersheds or bays	Near field/Far-field/Far-far field (optional) as defined by effluent concentration in receiving environment	MANOVA/ Multivariate
Simple Gradient (SG)	freshwater rivers or lakes, marine fjords, intertidal or narrow estuaries with no upstream flow	a series of reference stations with no or low effluent levels, situated towards the end of a declining gradient of mill effluent	single gradient through "near"-field, "far"- field and "far-far"-field (optional) levels of effluent in the receiving environment	Regression/ Multivariate
Radial or Multiple Gradient (RMG)	freshwater lakes, marine/estuarine coastal sites with bi-directional flow	multiple series of reference stations with no or low effluent levels situated towards the end of several transects of declining levels of mill effluent	multiple gradients through "near"-field, "far"- field levels of effluent in the receiving environment	Regression/ Multivariate

Figure 6.3 Control/impact design for freshwater rivers and streams.

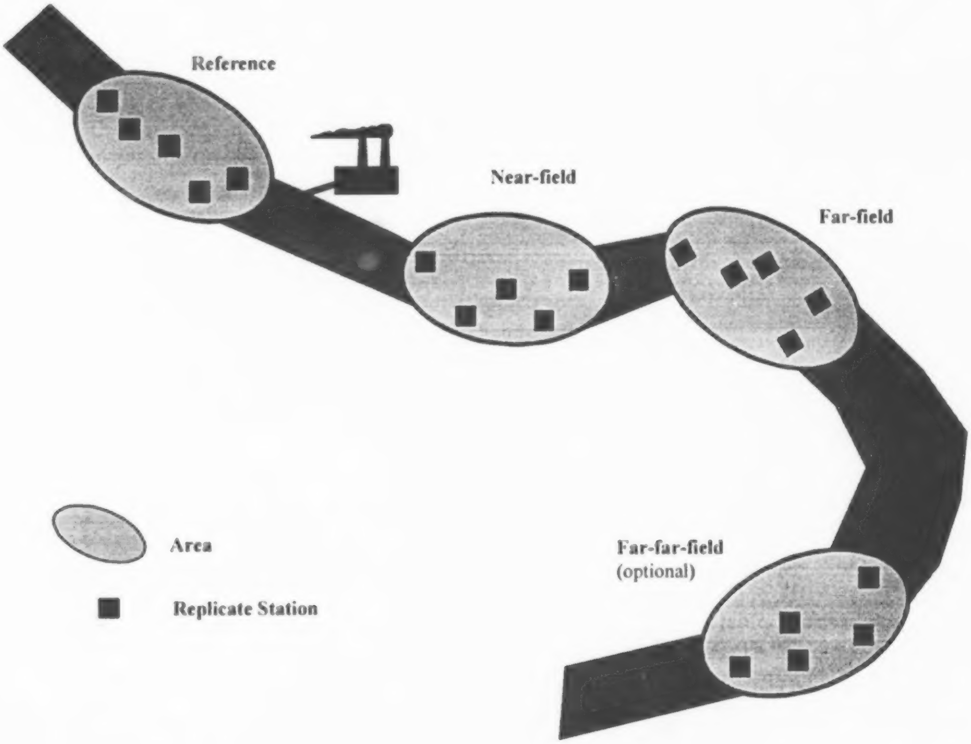
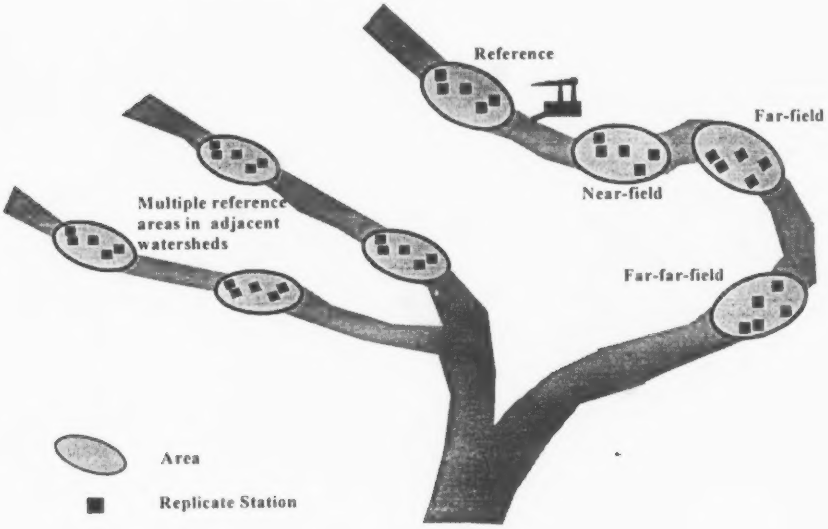

Reference
Near-field
Far-field
Area
Replicate Station
Far-far-field
(optional)

Figure 6.4 Multiple control/impact design

A: Freshwater rivers and streams


Reference
Far-field
Near-field
Multiple reference areas in adjacent watersheds
Far-far-field
Area
Replicate Station

B: Freshwater, spatially homogeneous lakes

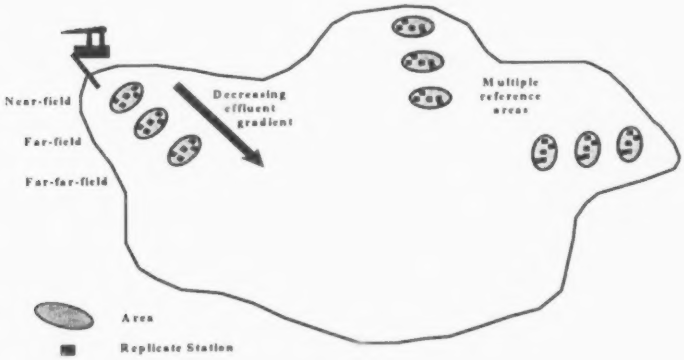

Near-field
Far-field
Far-far-field
Decreasing effluent gradient
Multiple reference areas
Area
Replicate Station

Figure 6.5a Linear gradient design for stream or river mills

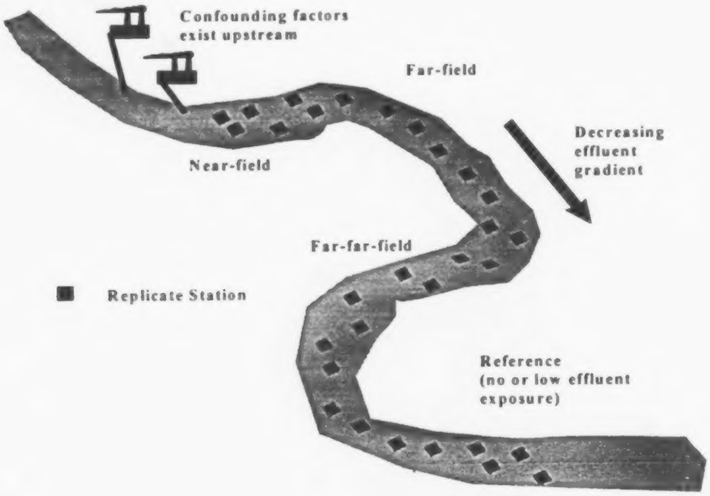

Confounding factors exist upstream
Far-field
Decreasing effluent gradient
Near-field
Far-far-field
Replicate Station
Reference (no or low effluent exposure)

Figure 6.5b Linear gradient design for lake or coastal mills situated in narrow bays or fjords.

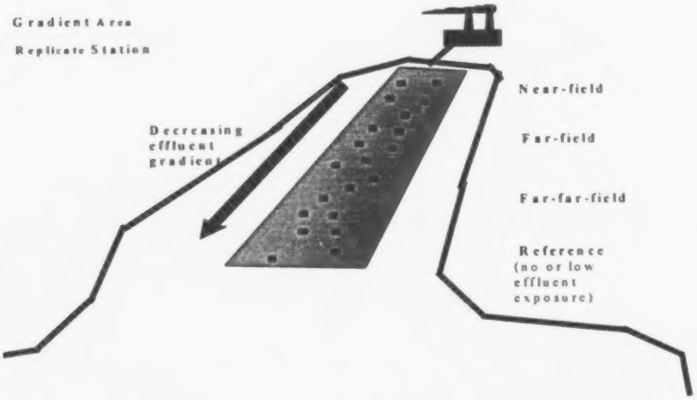

Gradient Area
Replicate Station
Decreasing effluent gradient
Near-field
Far-field
Far-far-field
Reference (no or low effluent exposure)

Figure 6.6 Radial gradient design for lake or coastal mills

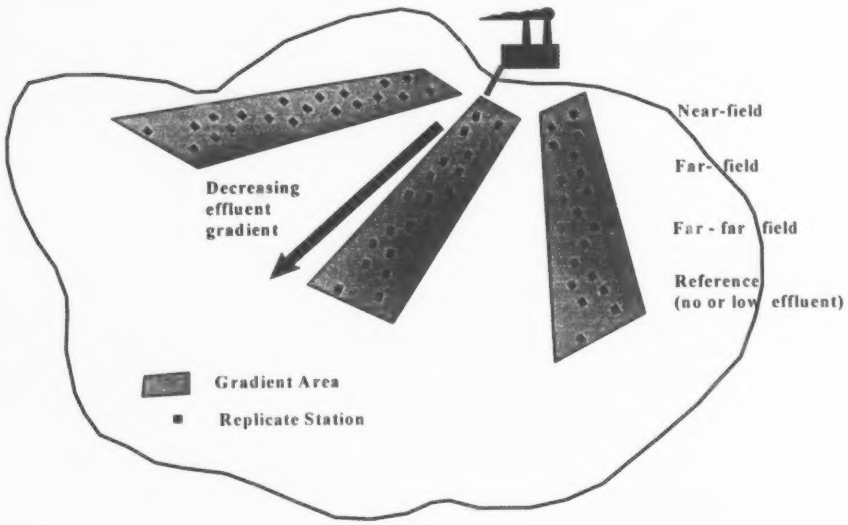
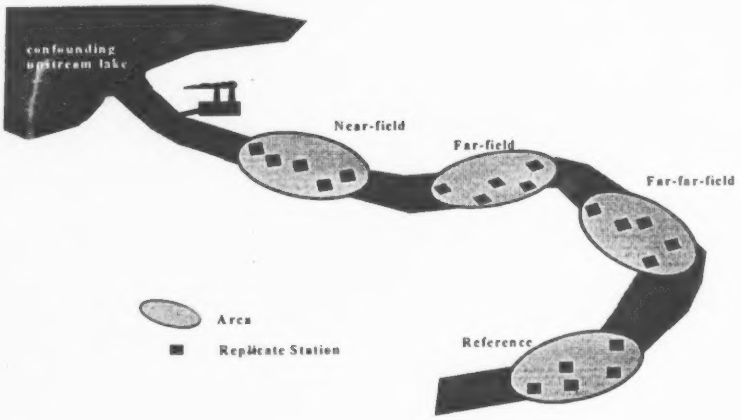

Near-field
Far-field
Far - far field
Reference
(no or low effluent)
Decreasing
effluent
gradient
Gradient Area
Replicate Station

Figure 6.7 Combination Design (C/I with downstream reference area) for river mills. Note that all mills are homogeneous in terms of environmental variables.


confounding
upstream lake
Near-field
Far-field
Far-far-field
Reference
Area
Replicate Station

6.1.5.5 Sampling Effort

Once the appropriate habitat and sampler types have been determined the amount of sampling at all levels must be addressed. In this section guidance is provide for determining the sampling effort required at both the **replicate station** and **field sub-sample** levels of scale.

Allocation of Replicate Stations

The number of replicate stations and their allocation between areas must first be determined such that the sampling design will answer the EEM monitoring requirements and questions for the specific site. Each area, will consist of a certain number of replicate stations, with the number of replicate stations to be determined through power analysis. The allocation and distribution of replicate stations is dependent on the site-specific conditions and the sampling pattern used (see Section 6.1.5.4) and guidance for calculating the number of replicate stations required to provide a given level of power in a Control/Impact ANOVA design is provided by Lowell (1997) and outlined below. Replication (i.e., the number of replicate stations) for gradient analysis designs should be no less than the sampling effort required for the basic C/I design.

For a complete example of a power analysis with Cycle 1 invertebrate community data please refer to Lowell 1997. An example of determining critical effect size from the variability inherent in reference areas is presented.

Determination of ES_{crit}

Mapstone (1995) gave a hypothetical example of the derivation of ES_{crit} by using the percent loss in coverage of a key algal species known to have serious, long-lasting effects on the functioning of that community. Unfortunately, ecological information at that level of detail is seldom available (Keough and Mapstone 1995). An alternative method of setting an ecologically relevant effect size is to define it in terms of the normal range of variability observed in relatively non-impacted reference communities; that is, to set ES_{crit} at the level of change where the mean for the impacted area is outside the normal range of values observed for unimpacted stations of the same habitat type in the reference area. A commonly accepted criterion for the normal reference condition is ±2 standard deviations (±2 SD) of the reference values about the reference mean, i.e., the 95% confidence interval about the reference area mean. Expressed as a percentage of the reference mean, this equals ±2 times the coefficient of variation for the reference mean (±2 CV).

This approach makes reasonable ecological sense and appears to have been applied successfully elsewhere (Kilgour 1997). For example, the 95% confidence interval for reference sites has been used to flag departure from normal reference conditions for freshwater zooplankton communities (Findlay and Kaisan 1996; Yan *et al.* 1996) and stream flow (Resh *et al.* 1988). The EEM Invertebrate Community Further Guidance document (Environment Canada and Department of Fisheries and Oceans, 1995a), in fact, also recommended using ±2 SD as an indicator of an important environmental effect, i.e., ES_{crit}.

Before ±2 SD can be calculated, a decision must be made on which field sub-samples to use. The more appropriate procedure is to calculate SD using variability among stations within the reference area(s) rather than variability among sub-samples within a station (e.g., sub-samples within a 10x10 m sampling station). The latter is at the wrong spatial scale to evaluate the normal range of variability for the reference area as a whole. The degree of variability among stations versus sub-sampling variability within stations will depend on the scale(s) of patchiness in invertebrate distribution, but using sub-sampling variability to test for area-wide effluent impacts effectively amounts to "pseudoreplication" (Hurlbert 1984). Consequently, although variability calculated from within station sub-samples would be valid for tests of differences among stations, the ANOVA contrasts for area-wide effects suggested in the EEM Invertebrate Community Further Guidance document (Environment Canada and Department of Fisheries and Oceans, 1995a) would be inappropriate.

Keough and Mapstone (1995) pointed out that stations (their term is "locations"), rather than sub-samples within stations, represent the effective experimental unit and can be considered to be random rather than fixed if they are taken to represent the status of the environment in general, irrespective of the number of stations that are physically available for sampling in a particular reference or impact area.

A key recommendation for subsequent cycles, is to increase the number of sampling stations within reference and impact areas, even if this necessitates fewer sub-samples per station. An increased number of sampling stations is required both to increase the accuracy of the among reference station variance estimates (used to calculate ES_{crit}) and to provide a large enough sample size to allow detection of a 2 SD difference between reference and impact means at a reasonable level of power (see Sample Size section below).

Determination of Sample Size (i.e., number of replicate stations)

Once generic ES_{crit} are set *a priori* for each habitat type, a recommended sample size can be determined. Of course, given unlimited resources, a very large sample size could reliably detect insignificantly small differences between the reference and impact area means (much smaller than a difference of ±2 SD). Very large sample sizes, however, are not economically feasible. Instead, a recommended minimum sample size can be calculated. The estimated sample size (number of replicate stations per reference and impact areas) required to detect a change of ±2 SD at a given level of power can be conveniently calculated by rearranging the standard power analysis equation as follows (Green 1989):

$$n = 2(t_\alpha + t_\beta)^2(SD/ES_{crit})^2$$

where:

n = sample size
t_α = t value for α significance level (Type I error; two-tailed)
t_β = t value for β significance level (Type II error; one-tailed)
SD = standard deviation
ES_{crit} = critical effect size

The power analysis equation can be modified by expressing ES_{crit} in terms of d, the number of standard deviations separating the reference and impact means ($d = ES_{crit}/SD$), and then substituting to yield:

$$n = 2(t_\alpha + t_\beta)^2/d^2$$

So, to determine the sample size to required to detect an ES_{crit} of 2 SD, d would be set equal to 2. The equation is solved iteratively by choosing an approximate value of n to look up t_α and t_β and then using the solution to find a more accurate n; the procedure is repeated until arriving at a final estimate for n. An alternative approach is to use the following equation, which yields an approximate n in one step (Guenther 1981; Alldredge 1987; Environment Canada and Department of Fisheries and Oceans 1995a):

$$n = (2(Z_\alpha + Z_\beta)^2/d^2) + 0.25(Z_\alpha)^2$$

where Z is the standard normal deviate for α or β. Pre-calculated tables of n are available for a variety of values of α, β, and d (Alldredge 1987; Cohen 1988; Environment Canada and Department of Fisheries and Oceans 1995a). Some relevant examples are given in Table 6.9.

Because of the potentially high cost (both ecological and monetary) of failing to detect negative impacts, many benthic impact researchers would argue that β should be set at least to the same level as α (Alldredge 1987; Underwood 1993; Mapstone 1995), which is, by convention, routinely set at 0.05. This is equivalent to setting power ($1-\beta$) at 0.95. Using these values and a ES_{crit} of 2 SD, a sample size of 8 stations per reference and impact area is recommended (Table 6.9). The minimum level of power normally used is 0.8 (Alldredge 1987; Cohen 1988; Burd *et al.* 1990; Osenberg *et al.* 1994; Keough and Mapstone 1995). At this level, a sample size of 5 stations per reference and impact area would be required (Table 6.9).

Table 6.9a Sample sizes required to detect difference of 1 SD

α	1-β			
	0.99	**0.95**	**0.90**	**0.80**
0.01	50	37	31	25
0.05	38	27	22	17
0.10	33	23	18	14

Table 6.9b Sample sizes required to detect difference of 2 SD

α	1-β			
	0.99	**0.95**	**0.90**	**0.80**
0.01	14	11	10	8
0.05	11	8	7	5
0.10	9	7	5	4

SD = standard deviation
See text for discussion of α, β, and 1-β.

Allocation of Field Sub-samples within a Replicate Station

In addition to determination of the total number of replicate stations through power analyses, the total number of field sub-samples to be taken at each replicate station must be considered. The issue of power for hypothesis testing and sampling precision are inter-related. For example, if 5 replicates are considered sufficient to obtain the power required for ANOVA type hypothesis testing, the question remains; do each of the 5 replicates accurately reflect the biotic descriptor in the sampled habitat? If not, then the ANOVA results may be unreliable. Therefore, the objective of multiple field sub-samples at each replicate station is to ensure that the sampling effort will produce an accurate reflection of all the metrics of interest (i.e. taxa richness, abundance, biomass) for the invertebrate community. The allocation of field sub-samples within a replicate station depends on two inter-related factors:

1) The abundance and degree of aggregation of organisms in relation to the desired level of precision for replicate station estimates.

For a given replicate station, the number of field sub-samples needs to be sufficient enough to give a mean and variance which provides confidence that a representative number of animals has been captured (for review see Burd *et al.* 1990). The more aggregated a community, the higher the variance of mean abundance for each replicate station. Elliott (1977) and Holme and McIntyre (1984) suggested the same simple method of determining the number of field sub-samples required to obtain a predefined level of precision. Elliott (1977) suggests that toleration of an index of precision (D) of 20% (i.e., that the standard error is equal to 20% of the mean) is acceptable for most bottom samples. The number of field sub-samples can then be calculated as follows:

$$n = s^2 / D^2 \ \bar{X}^2$$

where,

$\bar{X}$ = the sample mean
n = the number of field sub-samples
s^2 = the sample variance
D = the index of precision (i.e., 0.20)

Thus for the EEM program, to determine how many field sub-samples (i.e., grabs) per replicate station will provide an estimate with 20% precision, Cycle 1 data can be used to determine the mean and variance and hence, the appropriate number of field sub-samples. This determination may vary from location to location along with changes in the mean to variance ratio. It is recommended that the number of field sub-samples be calculated for locations which exhibit the highest variability and that the resultant sample size be applied equally to all areas to standardize sampling effort. Although this recommendation will produce better precision in the less variable habitats, it is a conservative approach and maintains equal sampling effort between areas and replicate stations. Also of note, is that with aggregated populations, although the overall mean should remain the same, depending on the scale of the aggregation in relation to sampler size, variance will change with the size of the sampler. Therefore, sample size estimates, using Cycle 1 or preliminary data are only relevant to a sampling program which would employ the same type and size of sampler with which the preliminary data was obtained. In cases where this sampling effort cannot be determined from a previous cycle's data, counting organisms in field sub-samples from the current survey as they are processed and calculating means and variances will allow determination of how many grabs should be processed in the laboratory. However, this *a posteriori* approach necessitates that a sufficient number of grabs was obtained during the field survey in the first place so the effort to calculate the sample size within a replicate station *a priori* should minimize problems with insufficient sampling effort.

Another approach which uses abundance and variance to determine sampling effort and precision but is more practical for determining the number of field sub-samples or replicates required at all replicate stations in a gradient-type design is given by the relationship between mean and abundance for all replicate stations in the gradient. Downing (1979, 1986) used Taylor's power law (1961) to estimate aggregation in a freshwater benthic community, and thus sampling effort required to reduce variance to an acceptable level. In effect, a given number of organisms is required for each replicate station in order to produce the required precision of mean abundance from sampling. Vezina (1988) used the same approach to determine empirically the degree of aggregation inherent in marine benthic communities. The power regression equation which describes the log/log "relationship" between mean abundance and variance for all locations, provides an formula which is then used to determine the estimated variance expected for a given abundance of organisms in that survey region. From this, the estimated variance for each mean abundance at each replicate station is calculated in the same way as Elliott (1977) to estimate the number of field sub-samples required at that replicate station. The only difference between the two methods of Elliott (1977) and Downing (1979) is that the variance used in the equation to determine replicate number is based respectively on; 1) sample variance and; 2) variance calculated for the sample mean from the power equation for all the samples in the survey region. Furthermore, the index of aggregation (slope) from the power regression equation can then be used to determine the most appropriate data transformation for parametric statistical analyses (see section 6.3.2 Data Transformations). Unfortunately, this method assumes that the overall assemblage has a uniform aggregation throughout the study area, which may or may not be true when an external environmental stressor is applied. However, the degree of goodness of fit of the mean and variance data to the log regression equation provides a good indication of how true the homogeneous aggregation assumption is. If there are extreme outliers (ie at replicate stations right outside the discharge point where abundance may be zero), they should be taken out of the analysis to avoid skewing the results. Because the aggregation of benthic communities can change as environmental conditions change either naturally or unnaturally, it is wise to review the relationship between mean and variance every time benthic samples are collected.

2) The number and distribution of different species in relation to obtaining a representative collection.

Determination that sufficient species have been sampled can be done using simple rarefaction methods such as the "species abundance curve" or species/ sampling area curve (for review see Burd *et al.* 1990) which compare the number of species obtained versus number of individuals for different numbers of pooled replicates. Using these two approaches, the total number of organisms that should be obtained for each replicate station can be determined. Because of assumptions inherent in the underlying distribution of fauna related to logarithmic species abundance curves, a more sophisticated approach is the "similarity/ sampling area" curve, which uses similarity indices on presence/absence data to determine the sampling effort required to obtain an acceptable overall faunal similarity between replicate stations (Weinberg 1978, Kronberg 1987).

If preliminary or cycle 1 data are unavailable or unsuitable for determination of the number of field sub-samples required to obtain a representative collection of species, a check on sampling effort could be performed very simply. If it is estimated that X number of grabs per replicate station are sufficient to achieve a data quality objective of retrieving 95% of benthic species present at any replicate station, more grabs can be collected at a few selected replicate stations and analyzed. Determination of a taxa richness plateau from these extra samples determines whether the number of grabs were sufficient to achieve the 95% objective (using a species area curve) .

6.2 Benthic Invertebrate Sampling Methods

This section outlines acceptable methods for the sampling of invertebrate communities and the processing of samples and refers to literature which discuss these methods in more detail. A lack of standardization makes it difficult to carry out a national or regional assessment of the environmental performance of mills. It was recommended that consistent methods be used for all EEM invertebrate community surveys, including standardization for habitat types, sampling devices, mesh sizes and sorting protocols.

6.2.1 Field Methods

6.2.1.1 Field Sieves and Mesh Sizes

Benthic samples typically contain varying amounts of fine sediment and debris. To expedite transfer to sample containers, storage, and shipping, these samples should be reduced in the field by sieving. Field sieving should be done wherever possible, immediately after sample retrieval and before preservation as many organisms become fragile and brittle after preservation and handling damage should be minimized. Furthermore, storing large volumes of fine sediment with preservative is less efficient than sieving and then preserving. Not only do organisms at the bottom and in the middle of a densely packed sediment not get sufficiently preserved, considerably more toxic preservative (i.e., formalin) is needed. Formalin must also be disposed of in the appropriate manner, often at additional cost, and in the interests of environmental conservation, the quantity of formalin used should be minimized wherever possible. Various techniques for sieving are available, but most involve washing the sample with a sieve bucket device. For marine organisms, samples should be sieved with seawater rather than freshwater, since the osmotic shock of freshwater may cause cell bursting and gross distortion of the animals and where appropriate, field water used to sieve should be screened for ambient organisms with a mesh smaller than the required minimum screen size used for the study. In addition, extreme care should be taken during washing of samples to avoid breakage of specimens, which can greatly reduce taxonomic efficiency and cost-effectiveness. Methods have been described to reduce breakage, particularly in marine samples (Gray *et al.* 1992).

In freshwater, macroinvertebrates are defined as those retained by mesh sizes of 200-500 µm (Slack *et al.* 1973; Weber 1973; Wiederholm 1980; Suess 1982), although immature life stages of some taxa may be smaller and some adult life stages may be larger. In marine systems, macrobenthos are typically those retained by sieves with 500-1000 µm mesh (Reish 1959, Thiel 1975, Pearson 1975, Holme and McIntyre 1984, Gray *et al.* 1992). It is estimated that a 1.0 mm sieve will retain about 95% of the biomass of marine macrofauna (Reish 1959) while reducing the numbers of juvenile taxa and meiofauna present in samples which respond functionally differently to environmental perturbation than adult macrofauna (Schwinghamer 1981, 1983, Warwick 1986). All freshwater mills must adopt a 500 µm mesh for field and laboratory use. All marine studies should use a stacked set of and 1.0 and 0.5 mm screens in the field with the 0.5 mm samples being archived and processed only if appropriate. Note that these mesh sizes are applicable to all equipment used in the field and laboratory i.e., both the Nitex mesh on the benthic samplers and sieving apparatus. The cost savings realized from the reduced taxonomic effort, resulting from the large mesh sizes, should be applied to improving the survey design and sampling coverage.

In some circumstances it may be desirable for the field samples to be screened for smaller organisms. If this is the case for an ICS, the data for the recommended sieve sizes must be reported independently so that national comparison for the EEM program can be readily performed without significant data searching. The possibility of including smaller sieve sizes in a study needs to be discussed with the RAO but some reasons for these additional measures include the following: 1) if sampling has to be conducted at times other than when organisms are large and abundance and diversity are high (Rees (1984), Barber and Kevern (1974) and Jonasson (1955) provide information on seasonal effects of mesh size, 2) if historical benthic surveys for the system under investigation utilized smaller mesh sizes, or 3) if smaller forms such as meiobenthos or juveniles are important in a specific case to elucidate mill effects. However, if smaller forms are important, simply adding an additional sieve may not fulfill this function. The appropriateness of the sampling techniques should be assessed for smaller forms. For marine environments, Gray *et al.* (1992) noted that collection of meiobenthos is most appropriately performed with corer samplers which are not recommended sampling devices for the EEM program. Thus, before simply screening for smaller organisms, appropriate protocols should be implemented and discussed the local RAO.

In these aforementioned cases, it is highly recommended that a stack of screens be used which minimally have the mandatory sieve sizes (i.e., freshwater - 0.5 mm and marine - 1.0 mm) and then any other smaller sizes which are appropriate. This procedure simultaneously allows site-specific concerns to be addressed and fulfills EEM requirements by allowing for national or regional comparisons to be conducted on the standardized mesh sizes. Note, that sieving with only the finest scale sieve can be done in the field, as long as the appropriate fractionation of the sample with the stacked set of sieves is performed in the laboratory before processing, remembering that care must be exercised in any sieving after preservation as per the initial discussion in this section.

6.2.1.2 Sampling Equipment

During the design phase of the survey, certain decisions need to be made concerning the selection of specific sampling equipment. The two major considerations for benthic surveys are mesh size (see Section 6.2.1.1) and quantitative sampling equipment. Quantitative sampling of benthic communities is carried out using devices that sample a known area or volume of habitat, such as grab samplers or stream net samplers. Each sampling device should be non-selective and suitable for a particular condition of substrate. Benthos collected from natural substrates provide an indication of historical and current stressors. Therefore, samplers that collect benthos communities from the bottom sediments are required unless this is not possible due to physical constraints. Different samplers may be used in the same survey if they are used to sample different habitat classes. Samplers are to be consistent within a habitat class between all stations and areas. For example, if both erosional and depositional habitat classes are to be sampled downstream of a mill outfall, it would be reasonable to use one of the recommended grabs for the depositional habitat, but a Hess-type sampler, for the erosional habitat. It is recommended that grab samplers with screens on the top and top-opening gates be used so that the bow-wave ahead of them is reduced and less substrate is lost, and for examination (and possibly sediment chemistry analysis) of undisturbed surface layer in sediment samples. Samplers that collect drifting or colonizing benthos are not currently approved because they are selective and are not comparable with EEM surveys from other mills. Also, the use of artificial substrates should be avoided unless there are no other alternatives available and they are approved by the RAO. Artificial substrates require at least two visits (i.e. installation and recovery) to obtain one sample and, unless they mimic the natural substrate, attract an uncharacteristic assemblage of invertebrates which do not integrate temporal responses to effluent effects.

Included in the original EEM Technical Guidance Document (Environment Canada & Department of Fisheries and Oceans 1993) was a review of the relevant quantitative samplers which may be used for collecting benthos as part of the EEM program. The review covers the types of habitat that each sampler can be used for, the types of organisms likely to be collected or missed, and the advantages and limitations of each type of sampler. Further standardization of benthic samplers is necessary for regional or national comparisons of ICS data to be valid. Specific recommendations were forwarded as to the appropriate samplers for the various habitat classes encountered during EEM ICSs. These options are discussed below as they pertain to each potential receiving environment and summaries of the advantages and disadvantages of the recommended samplers are presented in Table 6.10. The selection of a sampler may also be influenced by the type of sampler(s) used in previous surveys of that system. To ensure that surveys during this EEM cycle can be compared with previous historical surveys, it would be useful to use similar sampling equipment However, use of sampling devices other than those recommended must be approved by the RAO. For more detailed information please refer to the bibliographies on quantitative samplers and appropriate methodologies which have been prepared by Klemm *et al.* (1990), Eleftherious and Holme (1984), Elliott and Tullett (1978,1983) and Rosenberg (1978), Downing (1984) and Mason (1991a). See also Rabeni and Gibbs (1978) and Alberta Environment (1990).

Standardization of techniques does not only apply to the actual samplers but to the level of expertise required to correctly deploy the sampler. Employees using the sampler must be properly trained so that any areas where subjectivity may be interjected is limited. For example, when sampling erosional zones in rivers, the depth to which the substrate is disturbed within the net-sampler should be standardized since some individuals may be more enthusiastic in regards to "stirring" up the substrate than others. It is essential that the study leader be well versed in benthic invertebrate sampling and conduct effective training sessions with the employees who will be performing the field sampling. If training is done effectively operator error can be eliminated (Reynoldson and Rosenberg 1996).

Depositional Habitats - Freshwater

Grab samplers, sometimes called dredges, are devices with spring-loaded or gravity-activated jaws that "bite" into unconsolidated substrates (sand, silt, mud, etc.) to enclose a defined surface area of the bottom. These devices are generally lowered on a line or cable from a survey vessel to the bottom, usually with the aid of a winch. The effectiveness of these samplers will depend on the type of substrate, the depth of penetration, the angle the device strikes the bottom, completeness of the closure of the jaws, the loss of material during retrieval, the effect of the "shock wave" and subsequent washout of organisms when the grab strikes the substrate

In depositional zones in freshwater rivers or lakes, Ponar or Ekman grabs are suggested as standard samplers for EEM invertebrate surveys. In Cycle 1, only the Atlantic region used the van Veen grab with regularity and if no particular reason exists for this use then the Ponar or Ekman should be adopted in all regions.

Depositional Habitats - Marine/Estuarine

Depositional habitats in marine environments can be sampled with the Smith-McIntyre grab, a modified Van Veen grab which are suitable and available in Canada. A good review of marine sampling methods is available in Eleftheriou and Holme (1984). However, in shallow subtidal areas where there is not enough water depth to allow the deployment of the larger grabs, the use of a smaller, or mini(petite)-Ponar grab can be substituted. This grab is deployable from small inflatable boats and can be retrieved by hand.

Intertidal soft substrates may be sampled using any device which demarcates at least $0.1m^2$ area. The soft substrate is then removed to a standard depth of 10 cm using whatever device is appropriate. Note that in general, the lowest intertidal level available for sampling is preferred because less harsh physical conditions promote higher species richness and abundance.

Erosional Habitats - Freshwater

Stream-Net Samplers are devices used for collecting benthic organisms in erosional riverine environments. They use meshes of various sizes to sieve organisms from water flowing through the mesh after disturbance of a known area of the substrate. It is recommended that erosional habitats in freshwater environments be sampled with the Neill-Hess cylinder type samplers which allow unit area (typically, $0.1 m^2$) estimates to be made. However, one draw back to cylinder samplers in streams include incompatibility with size of substrate. For example, in some large rivers, mean particle size may be too great for the Neill-Hess cylinder to effectively sample the benthos. In such cases, a recently tested U-net sampler (Scrimgeour *et al.* 1993) can provide area limited samples and can be adjusted accordingly to the size of the substrate. This sampler has been used successfully for a range of substrate sizes (Glozier 1989) and can sample either individual stones or a delimited area. Kick net samplers, although they don't provide an area delimited estimate, have been used widely in the United Kingdom, the U.S.A., Australia and Canada for large scale monitoring programs (Reynoldson *et al.* 1995)

Erosional Habitats - Marine/Estuarine

In marine/estuarine environments, large unconsolidated sediments such as gravel may be sampled with grab samplers. If not, hard substrates in erosional habitats (intertidal and subtidal) should be sampled using quadrats with a minimum size of $0.1 m^2$. However, some other quantitative techniques may be required to collect marine shellfish and other large species. These may include hand collections by divers, remote photographic techniques from defined surface areas (Eleftheriou and Holme 1984, Gray et. al. 1992), and collections from defined boundaries along transects. An outline of a marine sampling protocol is described in a series of Puget Sound Estuary Program Reports (Tetra Tech, Inc. 1986a, 1986b, 1987).

The intertidal zone should be sampled if the effluent plume impinges substantially on it. Determining the tidal level of greatest interest for examining mill impacts will involve logistical considerations. Basically, the lower in the intertidal the surveys can be conducted, the better, since less harsh conditions create less patchiness in flora and fauna (intersample variability). Coastal plants and animals in this habitat typically exhibit vertical distributions that reflect gradients in environmental parameters such as air exposure, temperature, salinity, light intensity, abrasion due to logs or ice and wave shock. These gradients should be considered in planning and undertaking biological surveys in the intertidal environment. Sampling protocols for this area will be somewhat different from those described in the earlier sections (for review see Gray et. al. 1992). Wherever possible, semi-quantitative surveys using quadrat areas of $0.1 m^2$ should be done. Determining the substrate or habitat type to be sampled is dependent on sampling limitations and the dominant habitat present (see Section 6.1.5.2 for discussion on dominant habitat selection) However, if it is a not feasible or ecologically sound to collect samples, then visual surveys may be required. The exclusion of an ICS and use of visual surveys should first be approved by the local RAO. If approved, a visual survey would include approaches such as recording and mapping (at a gross scale 5,000:1) the major biological features for assessment of gross changes in the biological community.

Table 6.10 Summary criteria for recommended samplers

Sampler	Characteristics
Freshwater - *Erosional Habitats in Rivers*	
Hess Sampler or **Neill Cylinder Sampler** **AREA : 0.1 m^2**	• **Habitats and substrates sampled:** shallow, flowing streams, with good current, rubble substrate, mud sand, gravel • **Effectiveness:** quantitative when used by experienced biologist, performance depends on current and substrate • **Advantages:** encloses area sampled, easily transported or constructed, samples a unit area, stable platform; can be used in weed beds. • **Limitations:** difficult to set in some substrate types, that is large rubble; cannot be used efficiently in still, slow flowing streams.
U-net samplers **AREA : 0.06 m^2 but any size can be specified during construction**	• **Habitats and substrates sampled:** shallow, flowing streams, with good current, rubble substrate, mud sand, gravel, and large cobble • **Effectiveness:** quantitative when used by experienced biologist, performance depends on current and substrate • **Advantages:** easily transported or constructed, samples a unit area, stable platform; size of frame can be constructed to adapt to various substrate types. • **Limitations:** cannot be used efficiently in still, slow flowing streams.
Kick net samplers **AREA : no set area but D-frame net is typically 30 cm wide**	• • **Habitats and substrates sampled**: shallow, flowing streams, with good current, rubble substrate, mud sand, gravel, and large cobble • **Effectiveness**: performance depends on current and substrate • **Advantages:** use easily learned, easily transported and constructed, • **Limitations**: not truly quantitative, does not sample a unit area, standardisation set by controlling the amount of time kicking
Freshwater - *Depositional Habitats in Rivers and Lakes -*	
Standard Ponar Grab **AREA : 0.05 m^2**	• **Habitats and substrates sampled:** freshwater lakes, rivers, estuaries, and reservoirs with hard and soft sediments such as clay, hard pan, sand, gravel, and muck; somewhat less efficient in softer sediments. • **Effectiveness:** not entirely adequate for deep burrowing organisms in soft sediments; very efficient for coarse sediments; collects both qualitative and quantitative samples • **Advantages:** better penetration than other grabs; side plates and top screens reduce washout, shock waves and substrate disturbance; best quantitative grab sampler for freshwater use. • **Limitations:** a heavy grab that requires use of a boat with winch and cable; stones, pebbles, and other debris can hold jaws open causing loss of sample.
Ekman Grab (Standard, Large and Extra-Large) **AREA : 0.02 m^2, 0.05 m^2 and 0.09 m^2,**	• **Habitats and substrates sampled:** freshwater rivers, lakes, and reservoirs where there is little current; soft sediments, e.g., muck and silt. • **Effectiveness:** efficient only in soft sediments but weights can be added for deeper penetration in fine sand; collects both qualitative and quantitative samples. • **Advantages:** easy to operate by hand, can be pushed into substrate in shallow water; hinged doors at top reduce washout, shock waves, and disturbance of the substrate; comes in a range of sizes. • **Limitations:** light weight so that jaw will not penetrate hard substrates; jaws often do not close completely due to blocking of jaws or failure of closing mechanism; inefficient in deep water or where current is moderate.

Table 6.10 Summary criteria for recommended samplers *(con't)*

Sampler	Characteristics
Marine/ Estuarine - *Erosional Habitats-*	
Quadrat samlers **AREA : 0.1 m^2 frames**	• **Habitats and substrates sampled:** large gravel or rocky bottom intertidal and subtidal (samples may be collected, photographed, mapped etc). • **Effectiveness:** depends on visual skills of field person, sampling often limited to larger organisms field personnel. • **Advantages:** For subtidal surveys, SCUBA transects with quadrats for counting megafauna are the only semi-quantitative method available. When depth are too deep for SCUBA, an ROV camera may replace the diver. • **Limitations:** for visual surveys, smaller organisms are underrepresented, can be logistically difficult in subtidal areas with more work and time in the filed than samples which can be brought back to the lab, semi-quantitative only.
Marine/ Estuarine - *Depositional Habitats -*	
Petite Ponar Grab **AREA : 0.02 m^2**	• **Habitats and substrates sampled:** freshwater lakes, rivers, and reservoirs and estuaries with moderately hard sediments such as sand, silt, and mud; will not penetrate clay; somewhat less efficient in soft sediments and coarse gravel. • **Effectiveness:** limited depth penetration; not useful in clay because of light weight, half recommended sample area size requires twice as many grabs. • **Advantages:** good penetration for such a small grab; can be operated by hand. • **Limitations:** jaws can be blocked by stones, sticks and other debris causing loss of part of the sample; not efficient in swiftly flowing water; smaller size makes it inefficient for sampling larger benthic organisms; not entirely adequate for deep burrowing organism in soft sediments
Standard Ponar Grab **AREA : 0.05 m^2**	• **Habitats and substrates sampled:** freshwater lakes, rivers, estuaries, and reservoirs with hard and soft sediments such as clay, hard pan, sand, gravel, and muck; somewhat less efficient in softer sediments. • **Effectiveness:** not entirely adequate for deep burrowing organisms in soft sediments; very efficient for coarse sediments; collects both qualitative and quantitative samples • **Advantages:** better penetration than other grabs; side plates and top screens reduce washout, shock waves and substrate disturbance; best quantitative grab sampler for freshwater use. • **Limitations:** a heavy grab that requires use of a boat with winch and cable; stones, pebbles, and other debris can hold jaws open causing loss of sample.
Smith-McIntyre Grab **AREA : 0.1 m^2**	• **Habitats and substrates sampled:** marine and estuaries; adaptable to large rivers, lakes, and reservoirs with sand, gravel, clay, and similar substrates. • **Effectiveness:** has been widely used for sampling in marine and estuarine habitats, penetration - 17 mm in silt and clay, 5 dm in coarse sand.. • **Advantages:** provides reasonably quantitative samples; trigger plates help penetrate the substrate, top screens and stabilizing frame reduces washout and ensures upright landing. • **Limitations:** very heavy, needs boat and power winch; spring-loaded jaws could be hazardous; inefficient for collecting deep burrowing organisms; jaws can be blocked by debris.

Table 6.10 Summary criteria for recommended samplers *(con't)*

Sampler	Characteristcs
Marine/ Estuarine	***- Depositional Habitats (con't)***
Van Veen Grab **AREA : 0.1 or 0.2 m^2**	• **Habitats and substrates sampled:** marine and estuaries with sand, gravel, mud, clay, and similar substrates; could be adapted to freshwater. • **Effectiveness:** penetrates to a depth of 5 to 7 cm. • **Advantages:** jaws close better than the Petersen grab; samples most types of sediments; comes in a range of sizes. • **Limitations:** a very heavy grab that requires a large boat and power winch; jaws may become blocked by debris such as rocks and sticks; not useful for deep burrowing organisms.
Quadrat to depth of 10 cm **AREA : 0.1 m^2**	• **Habitats and substrates sampled:** intertidal marine/estuarine soft substrates to fine gravel. • **Effectiveness:** very effective with care of field personnel. • **Advantages:** inexpensive, allows sampling of substrates not accessible with grab samples • **Limitations:** may not sample deep burrowing organisms in sandy sediments.

Source: Eleftherious and Holme (1984), Klemm *et al.* (1990), Scrimgeour *et al.* (1993)

6.2.1.3 Specimen Preservation

All samples should be fixed in the field in 10% buffered formalin to prevent damage to freshwater and marine worms and polychaetes. Because formalin is a carcinogen and an irritant to workers, gloves and eye protection are needed to protect against splashes, and should be considered mandatory safety equipment. Furthermore, non-breakable sample jars should be sealed with parafilm, doubled bagged for transport back to the laboratory facilities and adequately labeled.

6.2.1.4 QA/QC for Benthic Invertebrate Field Operations

An outline of the recommended QA/QC requirements for the field components of the Invertebrate Community Survey is presented below. Most of the following information has been taken from Environment Canada and Department of Fisheries and Oceans (1993) and Gibbons *et al.* (1993).

Field Sampling is the first stage of data collection. QA/QC procedures for the Invertebrate Survey should be outlined during the development of the study plan and should be followed precisely in order to maintain high quality data. Field SOP should specify sampling equipment and protocols appropriate to the study and while a QA/QC plan for field sampling can have many components, some of the main procedures are listed below.

1) all personnel involved in the field sampling should have appropriate training and experience with field equipment and objectives,

2) all safety measures must be identified, understood and adhered to,
3) collection equipment must be appropriate for the specific water body and selected invertebrate group, and must be checked frequently and maintained on a regular schedule,
4) there should be some criteria for acceptability of samples obtained and clear directions if acceptability guidelines fail (i.e. when to retake a sample, for example, for a marine 0.1 m^2 grab sample a penetration of 10 cm depth and > 4 litres of sediment would be considered an acceptable sample (Gray *et al.* 1992),
5) contamination during chemical sampling should be checked with trip blanks and equipment rinsates,
6) field sieving, if necessary, must be done as soon as possible after retrieval of samples,
7) samples must be stored in appropriate containers with appropriate preservative to prevent breakage and spoilage,
8) all sample containers must be appropriately labeled,
9) detailed field notes should be maintained in a bound notebook,
10) chain-of-custody forms and appropriate shipping and storage procedures should be used.

For general QA/QC procedures and some examples of field check lists refer to Chapter 4.

6.2.1.5 Supporting Environmental Variables

Additional environmental information collected during each cycle should allow investigators to determine if, and to what extent, natural environmental variables within the major habitat classes influence or are related to benthic community structure in reference and exposure areas. A detailed list of which variables are required in each receiving environment is presented in Table 6.11. These variables are to be measured simultaneously with the ICS and information regarding their sampling is located in Chapter 7. Not all factors will be applicable to all habitats types. For example, salinity may only be important in intertidal or estuarine habitats. Sediment particle sizes, sufides and TOC are only applicable in unconsolidated sediments.

Other good sources of information on selecting habitat descriptors for explaining patterns in lotic invertebrate communities are Wright *et al.* (1984), Corkum and Currie (1987), Corkum (1991) and Ormerod and Edwards (1987). For lentic communities, papers by Johnson *et al.* (1991), Reynoldson *et al.* (1995) and Bailey *et al.* (1996) provide lists of the most useful habitat descriptors. Several authors (Clarke 1993; Clarke and Ainsworth 1993; Hargrave *et al.* 1993, 1997; Warwick 1993; and Warwick and Clarke 1993) have described responses in invertebrate communities and demonstrated the ability to distinguish natural variation from anthropogenic stressors.

Table 6.11 Supporting variables for freshwater (F) and marine/estuarine (M) habitats to be measured concurrently with benthic faunal surveys.

Habitat (F, M)	Variables required at all sites	Variables required if site-specific conditions dictate	Justification
F, M	Dissolved Oxygen		Dissolved oxygen can be decreased by mill effluents due to BOD
F, M	Water Temperature		Mill effluents may cause increases in water temperature in receiving waters.
F, M	Sediment TOC or LOI		Effects may be related to inputs from organic material.
F, M	Bed structure - Particle size, Matrix: framework ratio, degree of embeddedness, texture, colour and thickness of layers in cores		Differences in physical structure of the habitat can influence invertebrate community structure.
F, M	Depth		Water depth has a major effect on invertebrate communities.
F, M	Latitude, longitude		Provides station location information.
F, M		Colour or Turbidity	Mills may discharge effluents which are coloured. This may reduce the light available for primary production.
F,M		Current velocity	Can provide information on equivalence of sampling stations.
M	Salinity		Changes in estuarine conditions may affect benthic communities
M	Sediment C/N ratio		C/N ratio is affected by terrestrial (wood) versus aquatic plant sources of marine deposition
M	Sediment Eh (redox)		Sediment redox provides an indication of the anoxic/oxic boundary in sediments
M	Sediment total sulfides		Sulfides in marine sediments indicate the extent and nature of microbial response to organic enrichment
M		Light/wave exposure	information on natural habitat factors
M		Tidal cycle	information of natural stressors in habitat

Table 6.11 Supporting variables for freshwater (F) and marine (M) habitats to be measured concurrently with benthic faunal surveys (con't)

Habitat (F, M)	Key Variables	Site-Specific Variables	Justification
F	pH		Provides information on water quality.
F	Conductivity		Provides information on water quality.
F	Soluble Reactive or Total Dissolved Phosphorus and Total Phosphorus		Phosphorus is often the limiting nutrient in freshwater. Mills may discharge P which could lead to nutrient enrichment.
F		Nitrate-Nitrite, Ammonia, and Total Nitrogen	Nitrogen is often a secondary limiting nutrient in freshwater.
F		Dissolved Organic and Particulate Carbon.	Carbon is a nutrient source for microbes. Mills discharge quantities of these carbon sources.
F		Sodium	Na may be a good tracer in water .
F		Algal biomass as chlorophyll or ash-free dry mass	Algae in the water column (lakes) or on benthic substrates provides a food source for higher trophic levels.
F		Current velocity	Can provide information on equivalence of sampling stations.
F		Bankfull channel width	Provides habitat structure information
F		Alkalinity	Provides information for water quality.

6.2.2 Laboratory Methods

6.2.2.1 Sample Sorting

The first step, in processing invertebrate samples in the laboratory is to place the preserved sample into a sieve (freshwater = 500 μm; marine 0.5 or 1.0 mm where appropriate) and rinse the preservative and the residual fine debris from the sample. While flushing the samples, care must be taken not to damage the organisms with too much water pressure. Breakage and fragmentation of organisms is particularly a problem with marine samples and makes identification problematic. Samples can be sorted in water immediately, providing they are transferred back to preservative at the end of the work day (i.e., after 8 hours). Alternatively, samples may be transferred to 80% ethanol after the washing process and prior to sorting and taxonomic work. Transfer to ethanol within 48 to 72 hours may be necessary for marine samples to avoid the breakage of delicate structures as the organisms become brittle. After sorting or identification, most macro-invertebrates should be stored in a solution of 70 to 80% ethanol, and 5% glycerin in vials or jars with air-tight lids. If screw-cap lids are used, they should be sealed with wax film, and will require periodic checking (once or twice annually) to replace evaporative losses. Since preservatives can have a variety of effects on the lengths and weights of invertebrates (Leuven *et al.* 1985; Howmiller 1972), it is essential that the exact preservation protocol for all steps be drafted *a priori* and that it be consistently applied between all cycles, areas, replicate stationss and habitats. Disposal of used preservatives must be performed in accordance to provincial hazardous waste regulations (e.g., BC Special Waste Regulations, Alberta Waste Control Regulation, Ontario Regulation # 347 etc.)

For samples which have large pieces of organic matter, the samples can be divided in the laboratory into appropriate size fractions to expedite the sorting process. The most commonly used fractions are coarse (> 1.00 mm) and fine (500 µm - 1.00 mm), which correspond to the divisions used to define coarse and fine particulate organic matter (CPOM and FPOM, respectively). If there are very large pieces it is sometimes beneficial to separate these from the rest of the sample with a 4.00 mm sieve. All fractions should then be sorted accordingly and the large numbers of organisms warrants it, the size fractions can be sub-sampled independently (see Section 6.2.2.7). Careful note taking is required for these more complex sorting procedures so that densities are calculated accurately. After the initial washing and fractionation of samples, the invertebrates should be sorted by trained technicians from the debris on a gridded tray or petri dish under a dissecting microscope at 10X to 20X magnification.

Various additional techniques exist which are designed to speed up the manual sorting required for benthic invertebrate samples. These techniques include flotation of organisms in high-density solutions (e.g. magnesium sulphate, sodium chloride, calcium chloride, sugar, D-mannitol) or bubbling to trap animals in the surface film and to promote separation from denser sediment particles. Heavier organisms such as molluscs will not float and must be removed directly with forceps. Also, various stains such as rose bengal or phloxine B at a concentration of 100 mg/l can be used to make the organisms easier to see. A cautionary note is required, however, as the efficiency of many of these techniques differ between taxa. If these techniques are used, then the QA/QC performed on these samples should be designed to detect any differential sorting efficiency of taxa caused by the staining procedures.

Proper QA/QC procedures are critical at this stage and acceptable QA/QC guidelines for the processing samples for an ICS are outlined fully in section 6.2.2.6.

6.2.2.2 Laboratory Sub-sampling

Sub-sampling of the invertebrate sample in the laboratory is acceptable, providing that the method used is quantitative. Large samples or samples with large amounts of sediment debris may require laboratory sub-sampling prior to sorting. Various methods of laboratory sub-sampling are available, and are outlined in Klemm *et al.* (1990) as well as Canton (1991), Mason (1991b), Plafkin *et al.* (1989), Wrona *et al.* (1982), Marchant

(1989) and Sebastien *et al.* (1988). The method used should be thoroughly documented. Generally, a minimum of one-quarter of the sample should be sorted, however, a more definitive guideline is that sub-sampling should continue until a minimum number of animals is counted to obtain a predefined variance level. For example, for the cone subsampler described by Wrona *et al.* (1982), as the number of animals counted exceeds 50, the improvement in the error term for a given unit of effort decreases. The extra information gained by taking more sub-samples is not worth the extra time invested. In this case (Wrona *et al.* 1982) and others (Hickley 1975, Elliott 1977) the minimum number of animals counted was set at 100. It is recommended that this minimum number be determined for each sorting and laboratory sub-sampling protocol utilized during an EEM ICS using a similar approach.

Further rules or guidance for laboratory sub-sampling protocols include the following:

1) If a sample contains large numbers of a single taxa (e.g. >300 oligochaetes), or large numbers of organisms in a given fraction, sub-sampling can be performed on those fractions or for those organisms alone.
2) When undertaking a laboratory sub-sampling regime, sub-sampling error should be estimated by continuing to sort sub-samples until the entire sample is sorted for a minimum of 10% of all samples. This is especially important if laboratory sub-sampling is undertaken on more samples from one area than another.
3) Careful note taking is required for these laboratory sub-sampling procedures so that the accuracy of the calculations of total density and other metrics is not compromised.

6.2.2.3 Taxonomy

Examination of the effectiveness of higher (e.g. family) level taxonomic identifications in detecting the effects of pulp mill discharges will be conducted with the Cycle 1 and 2 data. Currently, however, the taxonomic evaluation of benthic samples should be carried out to the lowest taxonomic level readily possible - genus and species wherever feasible. Table 6.12 lists targeted levels of taxonomy desirable for major taxonomic groups of benthic organisms. In general, the level of taxonomy should be consistent in each major group for all samples from a survey, and from survey to survey. Organisms that cannot be identified to the desired level of taxonomic precision should be reported as a separate category in the fundamental dataset at the finest level of taxonomic resolution possible. Since the accuracy of the taxonomic work depends on the availability of up-to-date taxonomic literature, a basic library of identification keys is essential. Keys appropriate to the geographic region of study are recommended A detailed list of taxonomic references for various groups of freshwater organisms is provided by Klemm *et al.* (1990). Other key sources are found in Table 6.13. Microscope slide mounts should be prepared for taxa requiring detailed microscopic examination for identification. This may involve various steps including dissection, clearing, and staining. Slide preparation techniques are listed in Klemm *et al.* (1990). Organisms requiring slide preparation include oligochaetes and chironomids. For marine benthos, juvenile or non-adult fauna should be identified and enumerated separately from adults, as they show different patterns of response to environmental effects. All identifications should be carried out or verified by a qualified and experienced taxonomist.

6.2.2.4 Establishment of Reference Collections

For comparative purposes and quality control checks the maintenance of reference collections of organisms are strongly recommended . For each mill a complete reference collection with several (i.e., more than one (Green 1994) representative-sized specimens of each taxon should be compiled and archived. These collections should be updated each cycle as appropriate (i.e., with any new species or taxon which are newly collected) Each species should be verified by an external taxonomist and should encompass identified organisms from each area in the survey and the location of this collection reported. Photographic iconographs have been used to advantage (Camburn *et al.* 1984-1986). Reference collections are useful in confirming identifications, ensuring consistent taxonomy, and in training personnel. Protocols for reference and voucher collections for freshwater invertebrates are detailed in a report prepared for DOE's Fraser River Action Plan (Green 1994).

For marine samples it is suggested that these reference collections could be used for a further purpose; to develop a size and biomass database for each mill. For these purposed 5-10 representative specimens per species are recommended with mean width, lengths and blotted wet weights recorded for each group. This further information could be used for comparative purposes with historical data sets and/or to estimate gross changes in biomass due to enrichment.

Table 6.12 Recommended level of taxonomic precision for benthic macroinvertebrates

Freshwater Samples

Group	Level
Coelenterata	genus
Turbellaria	family
Nematoda	phylum
Nemertea	genus
Polychaeta	species
Oligochaeta:	
• Tubificidae	species (except some immatures)
• Naididae	species
• Lumbriculidae	species
• Other families	genus
Hirudinea	species
Bivalvia:	
• Pisidiidae	genus (*Pisidium*), species (*Sphaerium*)
• Unionidae	species
Gastropoda	species
Crustacea:	
• Ostracoda	order
• Harpacticoida	subclass
• Others	species
Acarina	order
Insecta	genus

Marine Samples

Group	Level
Porifera	class
Cnidaria	genus
Turbellaria	genus
Nemertea	genus
*Nematoda	phylum (not to be included in analyses)
Sipuncula	species
Brachiopoda	genus
Bryozoa	family
Mollusca	
• Aplacophora	genus
• Gastropoda	species
• Bia	species
• Polyplacophora	genus
• Scaphopoda	species
Annelida:	
• Polychaeta	species (except some immature)
• Oligochaeta	genus
Arthropoda:	
• Pycnogonida	family
• Cephalocarida	phylum
• Malacostraca	species
• *Copepoda	phylum (remove from analyses)
• Cirripedia	species
•	
Ascidiacea	family
Echinodermata	species

*nematodes and harpacticoid copepods are meiofauna andonly a fraction of specimens will be captured by a 0.5 mm of 1 mm screen - therefore numbers are not representative and should be excluded from analyses (Holme and McIntyre 1984)

Table 6.13 List of freshwater and marine taxonomic macroinvertebrate keys

Freshwater	Marine
Clifford 1991 (Alberta macroinvertebrates)	**Gosner 1971**
Lehmkuhl 1975a, 1975b, 1976, 1979 (Saskatchewan macroinvertebrates, mayflies and damselflies)	**Abbot 1974** (seashells)
Merritt and Cummins 1984	**Schultz 1969** (isopod crustaceans)
Stewart and Stark 1988 (stoneflies)	**Keen and Coan 1974** (Mollusca)
Clarke 1981 (molluscs)	**Austin 1985** (Pacific invertebrates)
Wiggins 1996 (caddisflies)	**Baker 1980** (Tubificid species)
Morihara and McCafferty 1979 (baetid mayflies)	**Banse 1972**, 1974 (polychaetes)
Edmunds *et al.* 1976 (mayflies)	**Appy et al. 1980** (Bay of Fundy polychaetes)
Peckarsky *et al.* 1990 (macroinvertebrates)	**Blake 1971** (Polydora, east coast)
Brinkhurst 1986 (Oligochaetes)	**Blake 1991** (Polychaeta, north Atlantic)
Klemm 1972, 1985, 1991 (leeches)	**Bousfield 1960** (Atlantic seashells)
Mackie et al. 1980 (Great Lakes molluscs)	**Brinkhurst 1982** (Oligochaetes)
Lewis 1974 (mayflies)	**Cutler 1973** (Sipuncula)
Morihara and McCafferty 1979 (Baetis larvae)	**Fournier and Petersen 1991** (Polychaeta)
Pennark 1978 (invertebrates)	**Laubitz 1972** (Caprellidaes)
Wiederhom 1983, 1986 (chironomids)	**Steele and Brunel 1968** (Amphipoda)
Bednarik and McCafferty 1979 (Stenonema)	**Wallace 1919** (Bay of Funday Isopoda)
Bousfield 1958 (amphipods)	**Kozloff 1987** (Pacific N.W. Invertebrates)
Burch 1975 a,b (Mollusca)	**Berkeley, 1952 a,b** (Pacific Annelida)
Chu and Cutkomp 1992 (immature insects)	**Bousfield and Hendryks 1994, 1995 a,b** (Pacific Amphipoda)
Fitzpatrick 1983 (Crustacea)	**Bousfield and Hoover 1995** (Pacific Amphipods)
Johannsen 1977 (Diptera)	**Bousfield and Kendall 1994** (Pacific Amphipods)
Oliver et al. 1978 (Diptera)	**Butler 1983** (Pacific shrimps)
Saether 1975 (Diptera)	**Coates 1980** (BC Enchytraeidae)
Saether 1977 (Chironimidae)	**Hart 1982** (BC crabs)
Schefter and Wiggins 1986 (Trichoptera)	**Hobson and Banse 1981** (BC polychaetes)
Schuster and Etnier 1978 (caddis fly)	**Lambert 1981** (BC sea stars)
Thorp and Covich 1991 (invertebrates)	**SBMNH 1994a,b,c, 1995a,b,c, 1996a,b,c, 1997a,b**
Wood 1963 (Diptera)	**Brunel 1970** (St. Lawrence benthos)
Bilyj 1991	**Meinkoth 1987** (seashore invertebrates)
Brandlova 1972 (Cladocera)	**Morris 1951** (Atlantic and Gulf Coast inverts)
Delorme 1991	**Theroux and Wigley 1983** (East Coast bivalves)
Dussart 1969 (copepods)	**Squires 1990** (Atlantic decapods)
Laplante et al. 1981 (Quebec coleoptera)	**Brunel 1960** (Atlantic Mysidacea)
Simpson and Bode 1990 (Diptera)	**Treadwell 1948** (Atlantic Annelida)
Walker 1933, 1955, 1958 (Odonata)	**Turgeon et al. 1988** (molluscs)

6.2.2.5 Invertebrate Biomass

Biomass provides additional useful information because it is related to the availability of energy to other trophic levels (i.e. fish). For marine communities some investigators suggest that an analysis of benthic abundance and biomass together, provide a sensitive indicator of changes in the composition of the benthic community (e.g. Warwick 1986; Warwick *et al.* 1987; Clarke 1990; Burd *et al.* 1990). Certainly, in cases where organic enrichment is a factor, biomass can be as important an indicator as species richness and abundance because it helps to illuminate the degree of utilization of organically enriched deposits by all trophic levels. For example, in marine samples, it is in the measurement of distributions of biomass that the three main functional groups of benthic organisms, microfauna (grain surface dwellers), meiofauna (interstitial organisms) and macrofauna (burrowers and epifauna) can be distinctly separated (Schwinghamer 1981, 1983). Because these three groups of organisms have different reproductive modes, metabolic rates, life-histories and habitat adaptations, they respond differently to habitat perturbation. Therefore it is important that benthic studies are not confounded by mixing these groups together in analyzing habitat effects. Thus, selection of mesh sizes is important (see section 6.2.1.1).

It is recommended that, for marine communities (not freshwater), some effort should be directed towards obtaining some measures of invertebrate biomass. However, because precise biomass measurements are time consuming and problematic (c.f. Crisp 1984) unless collected in more detail and more often than is feasible for EEM requirements, it is only possible to determine relative changes in biomass of samples for the EEM surveys. This can be simply done by taking blotted wet-weight measurements of representative sized adult specimens of each species for each survey. Since the method is non-destructive, the reference collection may be used for this purpose prior to external verification or archiving. The mean weight of a given species can then be used to transform species abundance data to relative species biomass data for further summary or statistical analyses. These data show relative, large scale changes only, and cannot be used to infer production or trophic flow rates within benthic communities.

6.2.2.6 QA/QC for Benthic Invertebrate Laboratory Operations

In the laboratory, invertebrate samples are processed and counts of the various taxa are made. The sorted, preserved samples from each survey should be retained in an appropriate storage facility for at least 5 years, or until it is determined that no further information, will be required from the samples. It is important that all samples are processed in similar manners such that the amount of experimental error in the counts is minimized. To minimize processing error, the following items should be included in the QA/QC program:

1) all personnel involved in the sample processing and analyses should have appropriate education and/or training,
2) the effects of sub-sampling (if done) on abundance estimates should be examined on a minimum of 10% of the samples and the effects of sub-sampling on the sample estimates be documented,
3) re-sorting of randomly selected samples should be done to determine the success of initial sorting (see details discussions below)
4) appropriate taxonomic references should be used given the type of habitat and geographic location,
5) a complete reference collection for each mill should be compiled and verified by an external taxonomic expert, and updated when appropriate (i.e., when new species or taxon are recorded)
6) a system for archiving samples should be outlined,
7) for marine and estuarine samples, taxonomic counts should separately enumerate and report adults and non-adults,
8) data handling, reductions, transformations and reporting methods should be indicated, and,
9) detailed sample processing and laboratory notes should be maintained in a bound notebook.

Ecological sample processing involves, as a first step, sorting of organisms from debris and, possibly, sub-sampling sorted organisms for detailed identification. Inevitably, processing errors are associated with these activities must be estimated (e.g. Kreis 1986, 1989). Verification of sorting efficiency is easily performed on a spot-check basis if the left-over debris from a sample is retained. It is recommended that at least 10 % of all samples be resorted, and that the criteria for an acceptable sort be that < 10% of the total number of organisms were missed. If > 10% of the total number is found during the resort, then all the samples within that group of samples requires resorting. A further criterion which would require a resort is if an entire group of benthic invertebrates was missed by the sorter, (i.e. it wasn't recognized as an organism), even if the missed organism constituted < 10 % of the total. The factors which should be considered when determining similar groups of samples include: 1) sampling area, 2) habitat class, and 3) individual sorters. The QA/QC guidelines apply independently to each group of samples sorted. Unsorted and sorted fractions are to be retained until taxonomy and sorting efficiency are confirmed and the data are reviewed by the RAO.

Verification of taxonomy by recognized experts is also recommended on a spot-check basis, and on any new specimen entering a reference collection. Museums are sometimes prepared to perform this service when remote areas are included in the study and new species or distribution records are likely. They will also sometimes serve as archives for selected samples or voucher specimens of organisms. However, since considerable effort is involved in long-term maintenance of preserved biological material, the quantity of material submitted should be minimized.

6.3 Data Analysis

Sampling design, data handling and data analyses should all be developed in concert, by a person or persons with adequate ecological and statistical expertise, and careful attention to answering the questions: What (if any) is the magnitude of the effects on benthic communities from the mill effluent? What is the spatial extent of measurable effects on benthic communities from the mill effluent? It should be quite clear that the recommended study designs (Section 6.x) are based on the assumption that the objective of an ICS is to detect differences between or among areas and that it is at this level the analyses should be performed. For guidance on critical effects, power determinations, confidence intervals, and weight of evidence approaches refer to chapter 3.

The data analysis methods recommended here are suggested as a minimum for reporting purposes and to provide some consistency among investigators. is recognized that there are a variety of methods for analyzing data describing benthic invertebrate communities. Therefore, the emphasis here is as much on graphical and tabular reporting of temporal and spatial trends in the communities, along with simple, documented, and accessible database creation and storage. If raw data are stored in a safe, accessible manner, then any re-analysis or meta-analysis considered appropriate can be done with minimal cost and maximum efficiency.

Some multivariate methods do not lend themselves to any type of statistical interpretation, but are graphical or illustrative in nature. It is particularly important not to "over-interpret" such results, or to carefully describe conclusions from these analyses while outlining both the ecological and mathematical assumptions and limitations of the method used. Nevertheless, these can be powerful tools for visualizing environmental effects. Results from these analyses would be used as a weight of evidence contribution to the conclusions for the study.

6.3.1 QA/QC for Data Input and Verification

The first step in any data analysis program is the proper verification of the data entry. Ecological community studies usually generate large data matrices when an observation on each taxon is recorded at every sampling station for hundreds of taxon. Computers and multivariate statistical programs generally assist in synthesis of the data to elucidate spatial and temporal patterns of community health (e.g. Hensler 1984; Ludwig and Reynolds 1988). Due to the nature and quantity of data, particular care is needed in data handling, to guard against data entry errors or subsequent misuse of data. Double entry systems and transcription checks against the original data records are useful QC techniques. Missing data should be clearly distinguished from a taxon absence by use of unique non-zero missing value codes with code definitions built into each file. Read-only files help to ensure data integrity. QA/QC concerns regarding data analysis include: data verification and validity, repeatability and robustness of statistical analyses, and rigour and defensibility of analysis. Gibbons *et al.* (1993) suggest that data analyses should be repeatable, such that if another investigator were to analyze the data set found in the Appendix of the report, they would arrive at the same conclusions using the same methods which were outlined in the report Other considerations regarding the data verification and analyses are listed below:

1) use trained and experienced personnel,
2) conduct screening studies to identify transcription errors, outliers and other suspicious data points,
3) provide raw data in a electronic database format and appendices to reports which summarize the data,
4) document the methods (specific statistical tests) and software (if used) used for analysis,
5) maintain detailed notes regarding the analyses of the survey data, and
6) present the results, interpretations, and conclusions clearly.

6.3.2 Data Handling

Dealing with outliers

The initial step in data analysis is to check for transcription errors. Failure to do this first invalidates further analyses. All computer entries must be verified by checking a hard copy of the file against the raw data sheets. This cross checking is preferably done by someone other than the original data entry person. Assuming the data are entered correctly, data should be summarized, screened for erroneous values and outliers, assessed for normality, and transformed if necessary (Gibbons *et al.* 1993).

Visual screening techniques such as Box and Whisker Plots, Normal Probability Plots and Stem-Leaf Diagrams can be used to identify extreme values (true outliers and/or data entry errors) (see Tukey 1977). Norris and Georges (1993) recommend examining abundance estimates for each taxon to determine if numbers are reasonable. They also recommend calculating means and standard deviations because aberrantly high or low values can indicate errors. Extreme values or outliers, which are not errors of some kind, should not be removed from the data set, because this will result in the loss of an observation and a loss of power to the EEM invertebrate survey. Instead, extreme values should be identified in the report and the influence of the extreme value(s) on the results should be determined by re-analyzing the data without the extreme value.

Unknown and Immature Organisms

Some samples may contain immature individuals that cannot be identified to the recommended level of taxonomic precision (Table 6.12). A similar situation could also occur when samples are improperly preserved and identifying features are destroyed (e.g. mollusc shells dissolve due to unbuffered formalin). For the purposes of correctly reporting the raw data, these unidentified taxa and their abundances should be provided within the electronic raw data and report appendices. However, for data analysis, investigators must decide whether or not to apportion the unknown individuals according to the ratio of known specimens. This assumes that the ratio of unidentified specimens is similar to the ratio identified specimens, which may or may not be true. The alternatives include: (1) not incorporating immature or damaged forms at all, (2) pooling all specimens (i.e., mature/immature, identified/unidentified) and lumping them into one category at the next higher taxonomic level or (3) keeping unidentified taxa as a separate category in the analysis . Option (1) is not preferred if the "problem" taxa represent a large proportion of the total benthic community. Option (2) assumes that all taxa within a higher taxonomic level respond the same way to effluent-related stressors, which may or may not be true. Option (3) will have variable effects on data interpretation depending on the abundance of unidentified taxa. Whatever choice is made will depend on the expertise and experience of the individual investigator, however, it should be fully documented in the methods section of the appropriate EEM reports.

For marine surveys, it is also required that immature and juveniles be counted and enumerated separately from adults whether or not they can be identified to species, so that the adult assemblage can be analyzed without the confounding influence form transient juveniles. Thus, data analyses should show results with and without immatures included. This is because newly settled benthic forms have different survival characteristics than adults, which have been present in the sediments much longer and integrate the effects of habitat perturbations over time. Depending on the timing of sampling, newly settled juveniles may be abundant in samples, but may all die within days due to habitat stressors, predation or competition. This is not to say that data on immatures is not important. Dramatic variations in immature settlement between nearby samples within physically homogeneous habitats may be indicative of varying levels of stressors. It is just important to avoid confounding of results by mixing the groups together for analysis.

Data Reductions

Data reductions should be done only to aid in statistical or multivariate analyses, and for the same reasons as data transformation (see below). Data reductions can include eliminating or rolling up rare taxa, or reducing field sub-samples by pooling or averaging. Protocols for data reductions for marine communities are varied, but subsequent interpretations of data analyses should take these reductions into account. For example, elimination of rare taxa may result in the elimination of 90% or more of the biomass within a given station if those rare taxa are large. In some cases, rare taxa are rolled up into higher groups, which prevents loss of information but adds assumptions about the uniform behavior of mixed taxonomic groups. Reviews of standard methods of data reductions are given in Stephenson and Cook (1980) with some ecological consideration in Burd *et al.* (1990).

Data transformations

Data transformations are often used without consideration as to the effects they have on interpretations of results. Transformations should be applied only with a complete understanding of their effect on the data and its interpretation and only if they are necessary to aid in statistical or multivariate analyses to:

1) normalize distributions,
2) reduce the effects of extremely dominant taxa within a data set on a multivariate summary (or ordination),
3) reduce the analytical problem of too many zeros in a data matrix (see Clarke and Green 1988),
4) make heterogeneous variances homogeneous or make the variance independent of the mean abundance for parametric analyses; and,
5) linearize relationships among variables.

Logarithmic transformations have often been used for benthic invertebrate data because organism abundance typically vary exponentially (Green 1979). A log transformation will reduce the importance of the numerically dominant members, and will improve the likelihood of resolving structure when differences are due to medium-abundance or rare taxa. However, a log transformation is quite extreme. Other researchers have advocated the use of other geometric conversions such as square-root, cube root, fourth-root, natural log, etc. (for reviews see Hoyle 1973, Tukey 1977 and Hoaglin *et al.* 1983, Downing 1981). Downing (1979) showed empirically that the best overall transformation for stabilizing variance in freshwater benthos was the fourth-root ($x^{0.25}$), because this greatly improves the performance of parametric multivariate methods such as ordinations. Vezina (1988) repeated the exercise for marine subtidal communities, concluding that they were empirically less aggregated than freshwater, requiring a less extreme transformation (e.g. $x^{0.4}$). However, both researchers emphasize that the mean and variance relationship of any given community must be analyzed to determine the most appropriate transformation. It is also possible in this way to check whether or not the transformation used has actually stabilized variance as required.

The most extreme form of transformation is a presence/absence conversion. This is simply a form of transformation that can be used to elicit structural differences giving all taxa equal importance. The scale of application of such a transformation should be clearly stated (e.g. sub-sample numbers transformed prior to averaging, or averages from sub-samples transformed).

Alternative approaches include data ranking or coding (i.e. 1 to 5 for abundance), data standardization (dividing values by the standard deviation either for the sample or the species), or relativization (using fractions of the maximum abundance sampled for a species over all replicate stations) (see also Burd *et al.* 1990).

It should be emphasized that the use of any transformation introduces assumptions which make the results of data analyses more complex to interpret. For example, assuming that the medium-rare taxa are more important in determining community structure that the dominant taxa in a stressed habitat is a large assumption. There is no a-priori ecological reason to make this assumption, unless the community is overwhelmed by "transients" such as newly settled juveniles which may or may not be able to survive the stressed conditions.

When there is strong dominance in faunal assemblages, a series of analyses with the most dominant taxa removed one at a time will provide some information about how the distribution patterns are affected by the medium-rare taxa. This approach falls into the category of "data exploration", which is one of the ecologist's most powerful suite of tools.

As an alternative to mathematical transformations, the use of a "biological" transformation such as biomass conversions for abundance data can be more readily interpreted ecologically, and will often have the same effect as geometric transformations (Burd 1992). In the absence of direct biomass measurements, a biomass conversion tends to de-emphasize the small, abundant forms and emphasize the large, rarer forms which are contributing considerably to the organic biomass of the community. Often, the comparison of both a raw-abundance and biomass-converted abundance analysis will be illuminating in discerning community structural patterns. Further comparison with traditionally geometric tranformed data analyses will help to determine how robust the patterns are, regardless of the method of data handling. The un-transformed data should be stored in accessible fashion to facilitate alternative transformations or analyses. Clarke and Green (1988) identify limitations of the log transformation, while Burd *et al.* (1990) provide a review of most commonly used transformations. Green (1979) discusses the merits of presence/absence data.

6.3.3 Statistical Analyses

The operative word here is "simplicity" as the message was very clear from Cycle 1 that the data-analysis requirements were overly complex and restrictive. The presence or absence, proportion, biomass or density of each benthic invertebrate taxon in the community is the basic information obtained in any benthic survey. Further recommendations of summary metrics to include in the fundamental data set are community descriptors such as taxonomic richness, diversity and equability and population descriptors such as mean and variance values of abundance and biomass. However, species richness and diversity indices are only useful if most of the fauna have been identified to species. If not identified to species level then only family, genera or other taxonomic group richness can be reported. Alternative classifications (e.g. sensitivity indices, functional group proportions or densities) can be used to re-express the data if sufficient independent information about ecological tolerances or preferences of community members is available.

It is recommended that communities be described at both the area and station spatial scales. For example, taxonomic richness can be expressed either as an average per area or as the total number of taxa observed at each replicate station. The same is true of other community descriptors listed above. Descriptions at both spatial scales are informative, especially in terms of potential meta-analysis of data from several EEM surveys.

The combination of a clear multivariate graphical description of the spatial pattern of community descriptors and simple statistical analyses such as one-way Analysis of Variance (ANOVA) or regression analysis will illustrate any site-specific effects of the effluent on the benthic invertebrate community. The recommended study designs from the Annex 1 lend themselves to a simple set of statistical analyses depending on the design. The basic analyses suggested for these study design are found in Table 7.8 but are repeated here:

Study Design	**Statistical Analysis**
Control/Impact	ANOVA/ Multivariate
Linear Gradient	Regression/ Multivariate
Multiple Control/ Impact	ANCOVA/ Multivariate
Radial Gradient	Regression/ Multivariate

Univariate

The standard approaches recommended for analyzing summary data (abundance, species richness, biomass, etc.) include univariate ANOVA (for use with Control/Impact) and regression analysis with gradient designs. The comparison of reference and (various degrees of) exposed areas with replication at the replicate station level, in the context of a single assessment study (ANOVA) is straightforward. In addition, power analyses and sampling design are straightforward (see Chapter 3).

For gradient type analyses, simple regression may be useful to compare relationships between univariate biotic factors such as abundance, biomass, species richness, abundance of major taxa, etc. and sediment physical and chemical factors. The regression design is more geared towards measuring residual variance for all replicate stations along the gradient in question. In other words, gradient analyses are less concerned with within and between-site variance (as in ANOVA) and more-concerned with degree of fit to a predicted relationship between biotic and environmental factors. For radial gradients, a comparison of regression patterns for each gradient (i.e. regressions of faunal abundance versus distance from the outfall) may help to illuminate the direction and extent of effects. Alternatively, all data from all gradients can be included in one regression, if the comparison is between biotic and physical factors unrelated to geographic or natural habitat factors (i.e. faunal abundance versus chlorophenols in sediments). In simple regressions, it is important to avoid confounding of the physical factor being used (i.e. sediment TOC may be confounded by sediment particle size). If sufficient sampling is done (i.e. radial gradients), it may be possible to pick and choose unconfounded replicate stations (i.e. homogeneous habitat conditions) to regress a biotic versus a mill-related variable.

If multiple physical or chemical sediment factors are being comprehend with a biotic metric, multiple regression analyses can be useful.

Multivariate

Because we cannot visualize patterns in multidimensional space (where each taxa represents a dimension in the assemblage distribution), it is necessary to reduce the dimensionality substantially to interpret trends. For community-level patterns, some form of multivariate or multi-dimensional analysis may be appropriate. Some of these examine the biotic community alone, whereas others combine comparative analyses of biotic and physical factors.

Similarity measures (for reviews see Clifford and Stephenson 1975, Washington 1984) may be used to determine the multi-dimensional (multi-species) differences in composition between replicate stations or samples. Such dimensional reductions are the raw material for multivariate graphical presentations such as ordinations and cluster analyses that some researchers feel are particularly valuable in identifying patterns in community variation among exposure areas. Ordinations are part of a family of parametric (linear) graphical statistical methods including MANOVA, discriminate function analyses, multiple regression, correspondence analyses, etc. which have underlying assumptions about normality and variance distributions in the community being analysed (see data tranformations above). Cluster analyses are also graphical, but do not generally have underlying linear assumptions about normality of distribution for the community. Neither method is inherently inferential or hypothesis-testing, however it is possible to overlay confidence limits on the mapped distributions. Methods and applications are available in multivariate statistics textbooks, and numerous computer packages are available.

There are also randomization, jackknife and bootstrap methods (distribution-free) which are non-parametric means of hypothesis testing (Efron 1982, Felsenstein 1985). Such a method has been applied to cluster analyses, to test the significance of homogeneity in groupings of replicate stations and separate distinct repeatable "assemblages" (Nemec and Brinkhurst 1988).

As with data transformations, it is important to understand the assumptions and limitations of any of these methods and how they affect the results of benthic community analyses, in order to aid in sound ecological interpretation of results.

Additional measures of data analysis, metrics, indexes or indicator species may be used if investigators have preferences or *a priori* reasoning that a further approach may be useful for a specific study. The intent of this guidance document is not to reproduce a statistical text with complete with error tables etc. Measures should be taken to ensure that analyses are appropriate to the study design and that other sources (Table 6.14) or a qualified biostatistician be consulted if necessary.

6.3.4 Reporting Requirements

It is recommended that once the data have been screened for coding or counting errors, all data should be stored in simple, accessible databases. This data base should be entered on a standard program (e.g. Microsoft Excel or Access) to facilitate comparisons of data sets among mills (but see Chapter 10). The format of this data base is currently being designed under an information management framework. Clearly, data need to be reported in a straight-forward, readily useable computerized format for larger-scale comparisons by regulators. The complete fundamental data set, including rare and highly variable taxa and ambiguous identifications should be stored in this manner, even if data filtering has been applied prior to calculation of community descriptors. This way, alternative approaches to data filtering, calculation of community descriptors, and analysis can be employed in re-analysis or meta-analysis. Specifically, a list of relevant details for the field, laboratory and data analysis components of an EEM invertebrate survey, which should be included with the report is provided below.

Field reporting requirements

1) field sheets are to be retained for five years,
2) replicate station location (grid coordinates),
3) date and time of sampling,
4) field crew members,
5) habitat descriptions including measures of the supporting environmental variables,
6) occurrence of mill-related solids and evidence of other contamination sources (discoloration, odor, petroleum residues) in sediments prior to processing samples,
7) sampling method used including type and size of sampler and sieve or mesh size,
8) photographs of replicate stations.

Laboratory reporting requirements

1) bench sheets are to be retained for five years,

2) raw data are to be reported for each replicate sample, listing taxa present, numbers of individuals,
3) degree of sorting efficiency achieved,
4) method and level of sub-sampling applied and sub-sampling precision
5) taxonomic authorities used,
6) location of reference collection and report on taxonomic verification.

Data analysis reporting requirements

1) tabular listing of the species abundance in each sample as an appendix,
2) tabular summaries of major taxon abundances (area means with standard errors),
3) tabular summaries of calculated univariate measures with variance estimates (e.g. abundance, richness, biomass),
4) a summary of the observed effect sizes, and,
5) estimates of power of the surveys based on the univariate measures above.,
6) a summary of adherence to data quality objectives, standard operating procedures and sampling protocols, and identification of any QA/QC problems,
7) effects of outliers or extreme values on the results (if any), and,
8) any other details relevant to determining the quality or interpretability of the data.

Table 6.14 Key references and sources

Benthic Invertebrate Ecology
Pearson and Rosenberg 1978
Schwinghamer 1981; 1983
Resh and Rosenberg 1984
Rosenberg and Resh 1993
Diaz and Schaffner 1990
Benthic Invertebrate Sampling and Sample Processing
Holme and McIntyre 1984
Downing 1979
Vezina 1988
Klemm *et al.* 1990
Stanford and Spacie 1994
Gray *et al.* 1990
Statistical Analyses
Green 1979, 1989
Burd *et al.* 1990
Elliott 1977
Clarke and Green 1988
Reynoldson *et al.* 1995

6.4 References

Abbott, R.T. 1974. American seashells. (2nd Ed.) Van Nostrand Reinhold Co., New York. 633 pp.

Alldredge, J.R. 1987. Sample size for monitoring of toxic chemical sites. Envir. Monitor. Assess. 9:143-154.

Alberta Environment. 1990. Selected Methods for the Monitoring of Benthic Invertebrates in Alberta Rivers. Environmental Quality Monitoring Branch, Alberta Environment.

Appy, T.D., I.F. Linkletter, and M.J. Dadswell. 1980. A guide to the marine flora and fauna of the Bay of Fundy: Annelida: Polychaeta. Fishereies and Marine Service, Technical Report No. 920. 124 pp.

Austin, W.C., 1985. An Annotated Checklist of Marine Invertebrates in the Cold Temperate Northeast Pacific, Vol. 1, 2 and 3, Khoyatan Marine Laboratory, Cowichan Bay, B.C., Contract to Fisheries and Oceans.

Bailey, R.C., Norris R.H., and T.B. Reynoldson. (eds.) 1996. Study design and data analysis in benthic macroinvertebrate assessments of freshwater ecosystems using a reference site approach. Technical Information Workshop, North American Benthological Society, 44th Annual Meeting, Kalispell, Montana. 69p.

Baker, H.R. 1980. Key to the Common Tubificid Species of the Northeast Pacific, Manuscript from Oligochaeta Workshop.

Banse, K. 1972. Redescription of some species of Chone Kroyer and Euchone Malmgren, and three new species (Polychaeta: Sabellidae) Fishery Bulletin 70(2):459-495.

Banse, K and K.D. Hobson, 1974. Benthic Errantiate Polychaetes of British Columbia and Washington, Bulletin of Fisheries Research Board of Canada, Bulletin 185, Ottawa.

Barber, W.E., and N.R. Kevern. 1974. Seasonal variation of sieving efficiency in a lotic habitat. Freshwat. Biol. 4:293-300.

Bednarik, A.F., and W.P. McCafferty. 1979. Biosystematic revision of the genus Stenonema. Can. Fish. Aquat. Sci. Bull. No. 201, 73 p.

Berkeley, C. and E. Berkeley. 1952a. Canadian Pacific Fauna, 9. Annelida 9b (1) Polychaeta Errantia, University of Toronto Press, Fish. Res. Bd., Canada.

Berkeley, C. and E. Berkeley. 1952b. Canadian Pacific Fauna 9. Annelida 9b (2) PolychaetaSedentaria, University of Toronto Press, Fish. Res. Bd., Canada.

Bilyj, B. 1991. Taxonomic verifications. Rapport interne. LRTAP Biomonitoring Program, Région du Québec. 13 p.

Blake, J.A. 1971. Revision of the genus Polydora from the east coast of North America (Polychaeta: Spionidae). Smithson. Contr. Zool. 75:1-32.

Blake, J.A. 1991. Revision of some genera and species of Cirratulidae (Polychaeta) from the western North Atlantic. Ophelia Suppl. 5:17-30.

Booth, J., Hay, D. and Truscott, J. 1996. Standard methods for sampling resources and habitats in coastal subtidal regions of British Columbia: Part I - Review of mapping with preliminary recommendations. Can. Tech. Rep. Fish. Aquat. Sci. 2118. 53pp.

Bousfield, E.L. 1958. Freshwater amphipod crustaceans of glaciated North America. Can. Field Nat. 72(2): April-June

Bousfield, E.L. 1960. Canadian Atlantic seashells. Canada Department of Northern Affairs and Natural Resources, National Museum of Canada. 72 pp.

Bousfield, E.L. and E.A. Hendryks, 1994. The Amphiphod Superfamily Leucothoidea on the Pacific Coast of North America. Family Pleustidae: Subfamily Pleustinae, Systematics and Biogeography, Amphipacific Vol.1, No. 2.

Bousfield, E.L. and E.A. Hendryks. 1995a. The Amphipod Family Pleustidae on the Pacific Coast of North America. Part III, Subfamilies Parapleustinae, Dactylopleustinae,and Pleusirinae: Systematics and Distributional Ecology, Amphipacifica Vol.II, No.1.

Bousfield, E.L. and E.A. Hendryks.1995b. The Amphipod Superfamily Eusiroidea in the North Pacific Region.I Eusiridae: systematics and distributional ecology, Amphipacifica Volume 1, Number 4.

Bousfield, E.L. and P.M. Hoover. 1995. The Amphipod Superfamily Pontoporeioidea on the Pacific Coast of North America. II. Family Haustoriidae. Genus Eohaustorias J.L. Barnard: Systematics and Distributional Ecology, Amphipacifica, Vol. II, No. 1.

Bousfield, E.L., and J. A. Kendall. 1994. The Amphipod Superfamily Dexaminoidea of the North American Pacific Coast; Families Atylidae and Dexaminidae: Systematics and Distributional Ecology, Amphipacifica Vol. 1, No. 3.

Brandlova, J. 1972. The cladocera of Ontario with remarks on some species and distribution. Can. J. Zool. 50: 1373-1403.

Brinkhurst, R. O. 1982. British and other marine and estuarine oligochaetes. Synopses of the British Fauna, N.S. 21: 1-127.

Brinkhurst, R.O. 1986. Guide to the freshwater aquatic microdrile Oligochaetes of North America. Can. Spec. Publ. Fish. Aquat. Sci. 84. 259 p.

Brunel, P. 1960. Artificial key to the mysidacea of the Canadian Atlantic continental shelf. Can. J. Zool. 38: 851-855.

Brunel, P. 1970. Catalogue d'invertébrés benthiques du golfe Saint-Laurent recueillis de 1951 à 1966 par la station de biologie marine de Grande-Rivière. Travaux sur les pêcheries du Québec, no. 32, 54 p.

Burch, J.B. 1975a. Freshwater Sphaeriacean clams (Mollusca: Pelecypoda) of North America, Malacological Publications, Hamburg, Michigan.

Burch, J.B., 1975b. Freshwater Unionacean Clams (Mollusca: Pelecypoda) of North America, Malacological Publications, Hamburg, Michigan

Burd, B.J., A. Nemec, and R.O. Brinkhurst. 1990. The development and application of analytical methods in benthic marine infaunal studies. Adv. Mar. Biol. 26:169-247.

Burd, B.J. 1992. Qualitative and quantitative studies of benthic infaunal communities in British Columbia coastal waters. PhD thesis, U. Victoria, Victoria, B.C.

Butler, T.H..1983. Shrimps of the Pacific Coast of Canada, Bulletin 202, Dept. of Fish. And Oceans, Ottawa.

Camburn, K.E., J.C. Kingston and D.F. Charles (eds.). 1984-1986. PIRLA Diatom Iconograph. Contains 53 photographic plates and 1059 figures, plus figure legends. PIRLA unpublished Report Series, Report No. 3. Indiana University, Bloomington, Indiana.

Canton, L.W. 1991. Improved subsampling methods for the EPA "Rapid Bioassessment" benthic protocols. Bull. N. Am. Ben. Soc. 8:317-319.

Chu, H.F. and L.K. Cutkomp, 1992. How to Know the Immature Insects, Wm. C. Brown Publishers, Dubuque, Iowa, 2nd edition.

Clarke, A. H. 1981. The freshwater molluscs of Canada. National Museum of Natural Sciences. Ottawa, Ontario. 446p.

Clarke, K.R. 1990. Comparison of dominance curves. Mar. Biol. Ecol. 138:143-157.

Clark, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. Aust.J.Ecol., 18:117-143.

Clarke, K.R. and M. Ainsworth. 1993. A method of linking multivariate community structure to environmental variables. Mar.Ecol.Prog.Ser., 92:205-219.

Clarke, K.R., and R.H. Green. 1988. Statistical design and analysis for a 'biological effects' study. Mar. Ecol. Prog. Ser. 46:213-226.

Clifford, H.T. and Stephenson, W. 1975. An introduction to numerical classification. Academic Press, London. 229 pp.

Clifford, H.F. 1991. Aquatic Invertebrates of Alberta. The University of Alberta Press. Edmonton, Alberta. 538p.

Coates, K. A.,1980. Keys to Intertidal Genera and species of Enchytraeidae found in British Columbia, Manuscript from Oligochaeta Workshop, U. Vic.

Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Lawrence Erlbaum Associates, Hillsdale, NJ, 567pp.

Corkum, L.D. and Currie, D.C. 1987. Distributional patterns of immature Simuliidae (Diptera) in northwestern North America. Freshwat.Biol., 17:201-221.

Corkum, L.D. 1991. Spatial patterns of macroinvertebrate distributions along rivers in eastern deciduous forest and grassland biomes. J.N.Am.Benthol.Soc., 10: 358-371.

Coull, B.C., and G.T. Chandler. 1992. Pollution and meiofauna: field, laboratory, and mesocosm studies. Oceanogr. Mar. Biol. Ann. Rev. 30:191-271.

Cowardin, L.M., V. Carter, G.C. Golet, and E.T. LaRoe. 1979. Classification of wetlands and deepwater habitats of the United States. U.S. Fish and Wildlife Service, FWS/OBS-79/31.

Cowardin, L.M. and Golet, C.F. 1995. US Fish and Wildlife Service 1979 Wetland Classification: a Review. Vegetatio 118: 139-152.

Crisp, D.J. 1984. Energy flow measurements, pp. 284-370 in Holme, N.A., McIntyre, A.D. (eds). Methods for the study of marine benthos. IBP Handbook 16. Blackwell Scientific Publications. 387 pp.

Cuffney, T.F., M.E. Gurtz, and M.R. Meador. 1993. Guidelines for the processing and quality assurance of benthic invertebrate samples collected as part of the national water-quality assessment program. U.S. Geological Survey Open-File Report 93-407.

Cutler, E.B. 1973. Sipuncula of the Western North Atlantic. Bull. Am. Mus. Nat. Hist. 152(3):1-204.

Diaz, R.J. 1992. Ecosystem assessment using estuarine and marine benthic community structure. In Sediment Toxicity Assessment (G. A. Burton, Jr.,ed),. Lewis Publishers, Chelsea, MI. pp 67-85.

Diaz, R.J., and L.C. Schaffner. 1990. The functional role of estuarine benthos. In Perspectives on the Chesapeake Bay, 1990. Advances in Estuarine Sciences (M. Haire and E.C. Krome, eds). Gloucester Pt., VA: Chesapeake Research Consortium, Rpt No. CBP/TRS 41/90, 1990, pp 25-56.

Delorme, L.D. 1991. Ostracoda, In: J.H. Thorpe & A.P. Covich (eds): Ecology and classification of North American Freshwater Invertebrates, Academic Press. 800 p.

Downing, J.A. 1979. Aggregation, transformation and the design of benthos sampling programs. J. Fish. Res. Bd Can. 36: 1454-1463.

Downing , J.A. 1981. How well does the fourth-root transformation work? Can. J. Fish. Aquat. Sci. 38: 127-129.

Downing, J.A. 1984. Sampling the benthos of standing waters. In A Manual on Methods for the Assessment of Secondary Productivity in Fresh Waters, Second Edition (J.A. Downing and F.H. Rigler, eds),. Blackwell Scientific, Oxford, pp. 87-130. IBP Handbook 17.

Downing, J.A. 1986. Spatial heterogeneity: Evolved behaviour or mathematical artefact? Nature, London 323: 255-257.

Dussart, B. 1969. Les copépodes des eaux continentales. N. Boubee & Cie (ed.) 292p.

Edmunds, G.F. Jr., S.L. Jensen, and L. Berner. 1976. The mayflies of North and Central America. University of Minnesota Press. Minneapolis. 330p.

Efron, B. 1982. The Jacknife, the Bootstrap and other resampling plans. SIAM, Philadelphia. 92 pp.

Eleftheriou, A. and Holme, N.A. 1984. Macrofauna techniques; pp. 140-216 in; Holme, N.A., McIntyre, A.D. (eds). Methods for the study of marine benthos. IBP Handbook 16. Blackwell Scientific Publications. 387 pp.

Elliott, J.M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. Freshwater Biological Assoc. Sci Publ. No. 25 160 pp.

Elliott, J.M., and P.A. Tullett. 1978. A bibliography of samplers for benthic invertebrates. Freshwater Biological Association, Occasional Publication, no. 4, 61 pp.

Elliott, J.M., and P.A. Tullett. 1983. A supplement to a bibliography of samplers for benthic macroinvertebrates. Freshwater Biological Association, Occasional Publication, no. 20, 27 pp.

Environment Canada & Department of Fisheries and Oceans. 1992. Aquatic environmental effects monitoring requirements. Annex I Aquatic environmental effects monitoring requirements at pulp and paper mills and off-site treatment facilities regulated under the Pulp and Paper Effluent Regulations of the *Fisheries Act* May 20, 1992. 23 (+iii) pp.

Environment Canada & Department of Fisheries and Oceans. 1993. Technical guidance document for aquatic environmental effects monitoring related to federal *Fisheries Act* requirements. Version 1.0. Environment Canada, April 1993, 128 pp.

Environment Canada, and Department of Fisheries and Oceans. 1995. Further guidance for the invertebrate community survey for aquatic environmental effects monitoring related to federal *Fisheries Act* requirements. Report No. EEM 2. Evaluation and Interpretation Branch, Environment Canada, Ottawa, ON, 206pp.

Environment Canada 1997a. Pulp and Paper Aquatic Environmental Effects Monitoring Requirements. ANNEX 1. EEM/1997/1. 52p.

Environment Canada 1997b. Benthic Invertebrate Community Expert Working Group Final Report. Recommendations from Cycle 1 Review. EEM/1997/7.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783-791.

Findlay, D. L. and S. E. M. Kaisan. 1996. The effect of incremental pH recovery on the Lake 223 phytoplankton community. Can. J. Fish. Aquat. Sci. 53:856-864.

Finlayson, C.M. and van der Valk, A.G. 1995. Wetland classification and inventory. A summary. Vegetatio 118: 1-2: 185-192.

Fitzpatrick, J. F. Jr., 1983. How to Know the Freshwater Crustacea, Wm. C. Brown Publishers, Dubuque, Iowa.

Fournier, J.A., and M.E. Petersen. 1991. *Cossura longocirrata* - Redescription and distribution, with notes on reproductive biology and a comparison of described species of Cossura (Polychaeta: Cossuridae). Ophelia Suppl. 5:63-80.

Frith, H.R., Seraring, G., Wainwright, P., Harper, H. and Emmett, B. 1993. Review of habitat classification systems and an assessment of their suitability to coastl B.C. Unpub. report to Environmental Emergency services Branch, Env. Can. from L.G.L. Ltd., Sidney, B.C.

Gibbons, W.N., M.D. Munn, and M.D. Paine. 1993. Guidelines for monitoring benthos in freshwater environments. Prepared for Environment Canada, North Vancouver, B.C. by EVS Consultants, North Vancouver, B.C. 81 pp.

Glozier, N. 1989. The effects of biotic and abiotic factors on the foraging success of a lotic minnow, Rhinichthys cataractae. M.Sc. Thesis, 151 p.

Gosner, K.I. 1971. Guide to the identification of marine and estuarine invertebrates. J. Wiley and Sons. New York. 693 pp.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley and Sons, Toronto. 257 pp.

Green, R.H. 1980. Multivariate approaches in ecology: the assessment of ecological similarity. Ann. Rev. Ecol. Syst. 11:1-14.

Green, R.H. 1989. Power analysis and practical strategies for environmental monitoring. Envir. Res. 50:195-205.

Green, G. 1994. Protocols for reference and voucher collections of aquatic invertebrates stored at the Royal British Columbia Museum. DOE-FRAP 1994-15.

Guenther, W. C. 1981. Sample size formulas for normal theory *T* tests. Am. Stat. 35:243-244.

Hargrave, B.T., Duplisea, D.E., Pfieffer, E. and D.J. Wildish. 1993. Seasonal changes in benthic fluxes of dissolved oxygen and ammonia associated with marine cultured Atlantic salmon. Mar.Ecol.Prog.Ser. 96: 249-257.

Hargrave, B.T., Phillips, G.A., Doucette, L.I., White, M.J., Milligan, T.G., Wildish, D.J. and R.E. Cranston. 1997. Assessing benthic impacts of organic enrichment from marine aquaculture. Wat.Air Soil Pollut. 99:641-650.

Hart, J.F.L.1982. Crabs and Their Relatives of British Columbia, Handbook No. 40, B.C. Provincial Museum.

Hay, D.E., Waters, R.D. and Boxwell (eds.) 1996. Proceedings Marine Ecosystem Monitoring Network Workshop, Nanaimo, B.C. March 28-30. 1995. Can. Tech. Rep. Fish. Aquat. Sci. 2108. 160pp.

Hensler, G.L. 1984. An introduction to statistical and ecological software. In: Statistics in the Environmental Sciences, Gertz, S.M. and M.D. London (eds.), STP 845, American Society for Testing and Materials, Philadelphia,

Hickley, P., 1975. An apparatus for subdividing benthos samples. Oikos 26:92-96.

Hoaglin, D.C., Mosteller, F. and Tukey, J.W. 1983. Understanding robust and exporatory data analysis. John Wiley, New York.

Hobson, K.D. and K. Banse. 1981. Sedentariate and Archiannelid Polychaetes of British Columbia and Washington, Bulletin 209, Ottawa.

Holme, N.A., McIntyre, A.D. (eds). 1984. Methods for the study of marine benthos. IBP Handbook 16. Blackwell Scientific Publications. 387 pp.

Howmiller, R.P. 1972. Effects of preservative on weights of some common macrobenthic invertebrates. Trans. Amer. Fish. Soc. 101:743-746.

Hoyle, M.H. 1973. Transformations - an introduction and a bibliography. Inter. Stat. Rev. 41:203-223.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. Pp. 31-47. In Biological Assessment and Criteria: tools for water resource planning and decision making (Davis, W.S. & T.P. Simon, eds.). Lewis Publishers, Boca Raton 415p.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monog. 54:187-211.

Johannsen, O.A., 1977. Aquatic Diptera, Entomological Reprint Specialists, LosAngeles, Calif., fourth reprinting.

Johnson, R.K., Wiederholm, T., and L. Eriksson. 1991. The influence of season on the classification and ordination of profundal communities of nutrient poor, oligo-mesohumic Swedish lakes using environmental data. Verh.Internat.Verein.Limnol.

Johnson, R.K., T. Wiederholm, and D.M. Rosenberg. 1993. Freshwater biomonitoring using individual organisms, populations, and species assemblages of benthic macroinvertebrates. In Freshwater Biomonitoring and Benthic Macroinvertebrates (D.M. Rosenberg, and V.H. Resh , eds),. pp 40-125.

Jonasson, P.M. 1955. The efficiency of sieving techniques for sampling freshwater bottom fauna. Oikos 6:183-207.

Keen, A.M. and E. Coan. 1974. Marine molluscan genera of western North America. Stanford University Press.

Keough, M.H. and B.D. Mapstone. 1995. Protocols for designing marine ecological monitoring programs associated with BEK mills. National Pulp Mills Research Program. Technical Report No. 11.

Kilgour, B.W. 1997. Ph.D. Thesis (in prep.). Biology Department, University of Waterloo, Waterloo, ON.

Klemm, D.J. 1972. Freshwater leeches (Annelida: Hirudinea) of North America. U.S. Environmental Protection Agency, Identification Manual no 8., 53 p.

Klemm, D.J. 1985. A guide to the Freshwater Annelida (Polychaeta, Naidid and Tubificid Oligochaeta, and Hirudinea) of North America, Kendall/Hunt Publ. Co. Dubuque, Iowa, 198 p.

Klemm, D.J. 1991. Taxonomy and pollution ecology of the Great Lakes region leeches (Annelida: Hirudinea). Michigan Academician. 24: 37-103.

Klemm, D.J., P.A. Lewis, F. Fulk, and J.M. Lazorchak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. EPA 600/4-90/030. U.S. Environmental Protection Agency, Environmental Monitoring Laboratory, Cincinatti, Ohio.

Kozloff, E.N. 1987. Marine Invertebrates of the Pacific Northwest, Univ. of Wash. Press.

Kreis, R.G., Jr. 1986. Variability study. Section 17. In Paleoecological Investigation of Recent Lake Acidification (PIRLA): Methods and project description (Charles, D.F. and D.R. Whitehead, eds). Electric Power Research Institute, Palo Alto, California.

Kreis, R.G., Jr. 1989. Variability Study - Interim Results. Section 4. In Paleoecological Investigation of Recent Lake Acidification (PIRLA): 1983-1985: Interim Report (Charles, D.F. and D.R. Whitehead, eds.). October 1989. Electric Power Research Institute, Palo Alto, California

Kronberg, I. 1987. Accuracy of species and abundance minimal areas determined by similarity area curves. Mar. Biol. 96: 555-561.

Lambert, P. 1981. The sea stars of British Columbia, B.C. Provincial Museum Handbook No. 39.

Laplante, S., Y. Bousquet, P. Bélanger, & C. Chantal. 1991. Liste des espèces de Coléoptères du Québec. Fabreries, Supplément 6. 136 p.

Laubitz, D.R. 1972. The Caprellidae (Crustacea, Amphipoda) of Atlantic and Arctic Canada. National Museum of Canada, Publications in Biological Oceanography, No. 4. 82 pp.

Lehmkuhl, D.M. 1975a. Field guide to aquatic insect families. Blue Jay 33:199-219.

Lehmkuhl, D.M. 1975b. Saskatchewan damselflies and dragonflies. Blue Jay 33:18-27.

Lehmkuhl, D.M. 1976. Mayflies. Blue Jay 34:70-81.

Lehmkuhl, D.M. 1979. How to know the aquatic insects. Wm. C. Brown Company Publishers. Dubuque, Iowa.168p.

Leopold, L.B., Wolman, G.M., and J.P. Miller. 1964. Fluvial processes in geomorphology. W.H. Freeman and Co., San Francisco. 522 p.

Leopold, L.B., 1994. A veiw of the river. Harvard University Press, Cambridge. 297 p.

Lewis, P.A. 1974. Taxonomy and ecology of Stenonema mayflies. EPA-670/4-74-006. 81 p.

Leuven, R.S.E.W., Brock, T.C.M., and H.A.M. van Druten. 1985. Effects of preservative on dry- and ash-free dry weight biomass of some common aquatic macro-invertebrates. Hydrobiologia 127:151-159.

Levings, C.D. and Thom, R.M. 1994. Habitat changes in the Georgia Basin: Implications for resource management and restoration. p 330-351 In Review of the marine environment and biota of Strait of Georgia, Puget Sound and Juan de Fuca Strait: Proceedings of the BC/Washington Symposium on the Marine Environment Jan 13,14 1994 (Wilson, R.C.H., Beamish, R.J., Aitkens, F. and Bell, J., eds). Can. Tech. Rep. Fish. Aquat. Sci. 1948.

Loeb, S.L., and A. Spacie. 1994. Biological Monitoring of Aquatic Systems. Lewis Publishers, Boca Raton, Florida. 381 pp.

Lowell , R.B. 1997. Discussion paper on critical effect size guidelines for EEM using benthic invertebrate communities. Report to the Environmental Effects Monitoring Program. EEM/1997/9

Ludwig, J.A. and J.F. Reynolds. 1988. Statistical Ecology: A Primer on Methods and Computing. John Wiley and Sons, New York.

Mackie, G.L., D.S. White, And T.W. Zdeba. 1980. A guide to freshwater molluscs of the Laurentian Great Lakes with special emphasis on the genus Pisidium. EPA-600/3-80-068. 143 p.

Malley, D.F., and J.B. Reynolds. 1979. Sampling strategies and life history of non-insectan freshwater invertebrates. J. Fish. Res. Bd. Can. 36:311-318.

Mapstone, B.D. 1995. Scalable decision criteria for environmental impact assessment: Effect size, Type I and Type II errors. In Detecting Ecological Impacts: Concepts and Applications in Coastal Habitats (Schmitt, R.J. and C.W. Osenberg, eds.), Academic Press, San Diego. pp. 67-80.

Marchant, R.A. 1989. A subsampler for samples of benthic invertebrates. Bull. Aust. Soc. Limnol. 12:49-52.

Mason, C.F. 1991a. Biology of Freshwater Pollution, Second Edition. Longman Scientific & Technical, New York. 351 pp.

Mason, W.T. Jr. 1991b. Sieve sample splitter for benthic invertebrates. J. Freshwat. Ecol. 6:445-449.

Matthews, G.W.T. 1993. The Ramsar Convention; its history and development. Ramsar Caonvention Bureau, Glan Switzerland

Meador, M.R., Hupp, C.R., Cuffney, T.F. and M.E. Gurtz 1993. Methods for characterizing stream habitat as part of the national water-quality assessment program. U.S. Geological Survey. Open-File Report 93-408, Raleigh, North Carolina.

Meinkoth, N.A. 1987. The Audubon Society field guide to North American seashore creatures. A.A. Knopf inc., 813 p.

Merritt, R.W., and K.W. Cummins. 1984. An Introduction to the Aquatic Insects of North America, Second Edition. Kendall/Hunt, Dubuque. 722 pp.

Morihara, D.K. and W.P. McCafferty. 1979. The Baetis larvae and North America (Ephemeroptera: Baetidae). Trans. Amer. Ent. Soc. 105:139-221.

Morris, P. A. 1951. A field guide to the shells of our Atlantic and Gulf coasts. The Peterson field guide series. Houghton Mifflin Company, Boston, Mass. 236 p.

Nemec, A.and Brinkhurst, R.O. 1988. Using the bootstrap to assess statistical significance in the cluster analysis of species abundance data. Can. J. Fish. Aquat. Sci. 45: 965-970.

Newbury, R.W. 1984. Hydrolic determinants of aquatic insect habitats. In The Ecology of Aquatic Invertebrates (V.H. Resh and D.M. Rosenberg, eds),. Praeger Publishers, New York.

Norris, R.H., and A. Georges. 1993. Analysis and interpretation of benthic macroinvertebrate surveys. In Freshwater Biomonitoring and Benthic Invertebrates (D.M. Rosenberg and V.H. Resh, eds.) pp. 234-286.

Oliver, D.R. et al, 1978. A Key to some Larvae of the Chironomidae (Diptera) from the Mackenzie and Porcupine River Watersheds, Fisheries and Marine ServiceTechnical Report 791, Agriculture Canada, Ottawa.

Ormerod, S.J. and R.W. Edwards. 1987. The ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye in relation to environmental factors. Freshwat.Biol., 17:533-546.

Osenberg, C. W., R. J. Schmitt, S. J. Holbrook, K. E. Abu-Saba, and A. R. Flegal. 1994. Detection of environmental impacts: natural variability, effect size, and power analysis. Ecological Applications 4:16-20.

Pearson, T.H. 1975. The benthic ecology of Loch Linnhe and Loch Eil, a sea-loch system on the west coast of Scotland. IV: Changes in the benthic fauna attributable to organic enrichment. J. Exp. Mar. Biol. Ecol. 20: 1-41.

Pearson, T.H. and Rosenberg, R. 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Ocean. Mar. Biol. Ann. Rev. 16: 229-311.

Peckarsky, B.L., Fraissinet, P.R., Penton, M.A., and Conklin, D.J. (eds), 1990. Freshwater macroinvertebrates of northeastern North America. Cornell University Press, 442p.

Pennak, R.W. 1978. Freshwater invertebrates of the United States: Protozoa to Mollusca. 3rd ed. John Wiley & Sons Inc., New York. 803 p.

Plafkin, J.L., M.T. Barbour, M.T. and K.D. Porter. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA/444/4-89-001, 162 pp.

Rabeni, C.F., and K.E. Gibbs. 1978. Comparison of two methods used by divers for sampling benthic invertebrates in deep rivers. J. Fish. Res. Bd. Can. 35:332-336.

Rees, H.L. 1984. A note on mesh selection and sampling efficiency in benthic studies. Mar. Poll. Bull. 15:225-229.

Rees, H.L., D.C. Moore, T.H. Pearson, M. Elliott, M. Service, J. Pomfret, and D. Johnson. 1990. Procedures for the monitoring of marine benthic communities at UK sewage sludge disposal sites. Department of Agriculture and Fisheries for Scotland. Scottish Fisheries Information Pamphlet Number 18, 1990, ISSN 0309 9105.

Reish, D.J. 1959. A discussion of the importance of the screen size in washing quantitative marine bottom samples. Ecology 40:307-309.

Resh, V.H., and D.M. Rosenberg. 1984. The Ecology of Aquatic Insects. Praeger Publishers.

Resh, V. H., A. V. Brown, A. P. Covich, M. E. Gurtz, H. W. Li, G. W. Minshall, S. R. Reice, A. L. Sheldon, J. B. Wallace, and R. C. Wissmar. 1988. The role of disturbance in stream ecology. J. N. Am. Benthol. Soc. 7:433-455.

Reynoldson, T. B., R. C. Bailey, K. E. Day, and R. H. Norris. 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology 20:198-219.

Reynoldson, T.B. and D.M. Rosenberg. 1996. Sampling strategies and practical considerations in building reference data bases for the prediction of invertebrate community structure. Pp 1-31. In Study design and data analysis in benthic macroinvertebrate assessments of freshwater ecosystems using a reference site approach (Bailey, R.C., Norris R.H., and T.B. Reynoldson. , eds.) Technical Information Workshop, North American Benthological Society, 44th Annual Meeting, Kalispell, Montana. 69p.

Robinson, C.L.K. and Levings, C.D. 1995. An overview of habitat classification systems, ecological models and geographic information systems applied to shallow foreshore marine habitats. Can. Man. Rep. Fish. Aquat. Sci. 2322: 65pp.

Robinson, C.L.K., Hay, D.E., Booth, J. and Truscott, J. 1996. Standard methods for sampling resources and habitats in coastal subtidal regions of British Columbia: Part 2 -

Review of sampling with preliminary recommendations. Can. Tech. Rep. Fish. Aquat. Sci. 2119. 119pp.

Rosenberg, D.M. 1978. Practical sampling of freshwater macrozoobenthos: a bibliography of useful texts, reviews, and recent papers. Canadian Fisheries and Marine Service Technical Report 790, 15 pp.

Rosenberg, D.M., and V.H. Resh (eds). 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman & Hall, New York. 488 pp.

Saether, Ole A., 1975. Nearctic and Palaearctic Heterotrissocladius (Diptera: Chironomidae_ Bulletin of the Fisheries Research Board of Canada, Bulletin 193, Ottawa.

Saether, Ole A., 1977. Taxonomic studies on Chironomidae: Nanocladius, Pseudochironomus and the Harnischia complex. Bulletin of the Fisheries Research Board of Canada, Bulletin 196, Ottawa.

Santa Barbara Museum of Natural History (SBMNH). 1994a. Taxonomic Atlas, Vol. 1, Platyhelminthes, and Nemertea, Santa Barbara, California, 1994

Santa Barbara Museum of Natural History (SBMNH). 1994b. Taxonomic Atlas, Vol. 2, Porifera.

Santa Barbara Museum of Natural History (SBMNH). 1994c. Taxonomic Atlas, Vol. 4, Oligochaeta and Polychaeta: Phyllodocidae to Paralacydoniidae.

Santa Barbara Museum of Natural History (SBMNH) 1995a. Taxonomic Atlas, Vol.5, The Annelida, Part 2 Polychaeta: Phyllodocida (Syllidae and Scale-Bearing Families), Amphinomida and Eunicida.

Santa Barbara Museum of Natural History (SBMNH). 1995b. Taxonomic Atlas, Vol. 12, The Crustacea, Part 3, The Amphipoda.

Santa Barbara Museum of Natural History (SBMNH). 1995c. Taxonomic Atlas, Vol. 13, The Bryozoa,

Santa Barbara Museum of Natural History (SBMNH). 1996a. Taxonomic Atlas, Vol. 6, The Annelida, Part 3 Polychaeta: Orbiniidae to Cossuridae.

Santa Barbara Museum of Natural History (SBMNH). 1996b. Taxonomic Atlas, Vol. 9, The Mollusca, Part 2 The Gastropoda.

Santa Barbara Museum of Natural History (SBMNH). 1996c. Taxonomic Atlas, Vol. 14, Miscellaneous Taxa.

Santa Barbara Museum of Natural History (SBMNH). 1997a. Taxonomic Atlas, Vol 10, The Pycnogonida, The Crustacea Part 1 The Decapoda and Mysidacea, 1997

Santa Barbara Museum of Natural History (SBMNH). 1997b. Taxonomic Atlas, Vol. 11, The Crustacea Part 2, The Isopoda, Cumacea and Tanaidacea.

Schefter P. and G.B. Wiggins, 1986. A Systematic Study of the Nearctic Larvae of the Hydropsyche morosa Group (Trichoptera: Hydropsychidae) Royal Ontario Museum, Toronto.

Schultz, G.A., 1969. The Marine Isopod Crustaceans, Wm. C. Brown Co., Dubuque, Ia.

Schuster, G.A. and D.A. Etnier, 1978. A Manual for the Identification of the larvae of the Caddisfly Genera Hydropsyche Pictet and Symphitopsyche Ulmer in Eastern and Central North America, U.S. Environmental Protection Agency, Cincinnati, Ohio.

Schwinghamer, P. 1981. Characteristic size distribution s of integral benthic communities. Can. J. Fish. Aquat. Sci. 38: 1255-1263.

Schwinghamer, P. 1983. Generating ecological hypotheses from biomass spectra using causal analysis: A benthic example. Mar. Ecol. Progr. Ser. 13: 151-166.

Scrimgeour, G.J., J.M. Culp and N.E. Glozier. 1993. An improved technique for sampling lotic invertebrates. Hydrobiologia 254:65-71

Scott, D.A. and Jones, T.A. 1995. Classification and inventory of wetlands. Vegetatio 118: 1-16.

Sebastien, R.J., Rosenberg, D.M. and A.P. Wiens. 1988. A method for subsampling unsorted benthic macroinvertebrates by weight. Hydrobiologia 157:69-75.

Simpson, K., & R. Bode. 1980. Common larvae of Chironomidae (Diptera) from New York state streams and rivers, with particular reference to the fauna of artificial substrates. Bull. New York state Mos. 439: 1-105.

Slack, K.V., R.C. Averett, P.E. Geeson, and R.G. Lipscomb. 1973. Methods for collection and analysis of aquatic biological and microbiological samples. In Techniques of Water-Resources Investigations of the United States Geological Survey, Chapter 4A, Book 5, pp. 1-165. U.S. Department of the Interior, Geological Survey, Washington, D.C.

Snelgrove, R.V.R. and C.A. Butman. 1994. Animal sediment relationships revisited: cause versus effect. Oceanography and Marine Biology annual review 32:111-178.

Squires, H.J. 1990. Decapod crustacea of the Atlantic coast of Canada. Department of Fisheries and Oceans. Ottawa. 532 p.

Stanford, L.L., and A. Spacie. 1994. Biological Monitoring of Aquatic Systems. Lewis Publishers, Boca Raton, Florida.

Steele, D.H., and P. Brunel. 1968. Amphipoda of the Atlantic and Arctic coasts of North America: Anonyx (Lysianassidae). J. Fish. Res. Bd. Can.. 25(5).

Stephenson, W. and Cook, S.D. 1980. Skewness of data in the analysis of species-in-sites-in-times. Proc. Royal Soc. Queensland 91:37-52.

Stewart, K.W. and B.P. Stark. 1993. Nymphs of North American Stonefly Genera (Plecoptera). University of North Texas Press, Denton, Texas. 460p

Suess, M.J., ed. 1982 Examination of Water For Pollution Control. A Reference Handbook. Vol. 3. Biological, Bacteriological and Virological Examination. Permagon Press, Oxford, England.

Taylor, L.R. 1961. Aggregation, variance and the mean. Nature (London) 189:732-735.

Tetra Tech, Inc. 1986a. General QA/QC Considerations for Collecting Environmental Samples in Puget Sound. A Report prepared for U.S. EPA, Region 10, Seattle. Final Report TC-3991-04.

Tetra Tech, Inc. 1986b. Recommended Protocols for Station Positioning in Samples in Puget Sound. A Report prepared for U.S. EPA, Region 10, Seattle. Final Report TC-3090-05.

Tetra Tech, Inc. 1987. Recommended Protocols for Sampling and Analyzing Subtidal Benthic Macroinvertebrate Assemblages in Puget Sound. A Report prepared for U.S. EPA, Region 10, Seattle. Final Report TC-3991-04.

Theroux, R.B. and R.L. Wigley. 1983. Distribution and abundance of east coast bivalve mollusks based on specimens in the National Marine Fisheries Service Woods Hole Collection. National Oceanic an Atmospheric Administration Technical Report (NMFS SSRF-768). 172 p.

Thiel, T.H. 1975. The size structure of the deep-sea benthos. Int. Rev. ges. Hydrobiol. 60:575-606.

Thorp, J.H. and A.P. Covich, 1991. Ecology and Classification of North American Freshwater Invertebrates, Academic Press, San Diego, New York, Toronto.

Thorson, G. 1957. Bottom communities(sublitorral or shallow shelf). In Treatise on Marine Ecology and Paleoecology.J.W. Hedgpeth, ed.Vol 1, pp. 461-534.Memoirs of the Geological Society of America 67.

Treadwell, A.L. 1948. Canadian Atlantic fauna 9. Annelida. 9b Polychaeta. Fish. Res. Board Can. 69 p.

Turgeon, D.D., A.E. Bogan, E.V. Coan, W.K. Emerson, W.G. Lyons, W.L. Pratt, C.F.E. Roper, A. Scheltema, F.G. Thompson and J.D. Williams. 1988. Common and scientific names of aquatic invertebrates from the United States and Canada: Mollusks. American Malacological Union. Bethesda, Maryland. 277 p.

Tukey, J. 1977. Exploratory data analysis. Addison-Wesley, Reading, Mass. 688pp.

Underwood, A. J. 1993. The mechanics of spatially replicated sampling programs to detect environmental impacts in a variable world. Australian Journal of Ecology 18:99-116.

Vezina, A.F. 1988. Sampling variance and the design of quantitative surveys of the marine benthos. Mar. Biol. 97: 151-155.

Walker, E.M. 1933. The nymphs of the Canadian species of *Ophiogomphus* Odonata, Gomphidae. Can. Ent. 65: 217-229.
Walker, E.M. 1953. The Odonata of Canada and Alaska. Part I, General, Part II. The Zygoptera-damselflies. Vol. I. Univ. Toronto Press, Toronto. 292 p.

Walker, E.M. 1958. The Odonata of Canada and Alaska. Anisoptera. Vol. 2. Univ. Toronto Press, Toronto. 318 p.

Wallace, N.A. 1919. The Isopoda of the Bay of Fundy. Univ. of Toronto Studies, Biol. Series No. 18. 42 pp.

Warwick, R.M. 1986. A new method for detecting pollution effects on marine benthic communities. Mar. Biol. 92: 557-562.

Warwick, R.M., and K.R. Clarke. 1993. Increased variability as a symptom of stress in marine communities. J. Exp. Mar. Biol. Ecol. 172:215-226.

Warwick, R.M. 1993. Environmental impact studies on marine communities: Pragmatical considerations. Aust.J.Ecol., 18:63-80.

Warwick. R.M., Pearson, T.H., and Ruswahyuni. 1987. Detection of pollution effects on marine macrobenthos: Further evaluation of the species abundance/biomass method. Mar. Biol. 95: 193-200.

Washginton, H.G. 1984. Diversity, biotic and similarity indices: A review with special relevance to aquatic ecosystems. Water Res. 18: 653-694.

Weber, C.I., ed. 1973. Biological Field and Laboratory Methods for Measuring the Quality of Surface Waters and Effluents. EPA-670/4-73-001. U.S. Environmental Protection Agency, Cincinnati, OH.

Weinberg, S. 1978. The minimal area problem in invertebrate communities of Mediterranean rocky substrata. Mar. Biol. 49: 33-40.

Wiederholm, T. 1980. Use of benthos in lake monitoring. J. Wat. Poll. Cont. Fed. 52:537-547.

Wiederhom, T. (ed.) 1983. Chironomidae of the holarctic region. Keys and diagnosis. Part 1 - Larvae. Ent. Scand. Suppl. No. 19, 457 p.

Wiederholm, T. (ed.) 1986. Chironomidae of the holarctic region. Keys and diagnosis. Part 2 - Pupae. Ent. Scand. Suppl. No. 28, 482 p.

Wiggins, G.B. 1996. Larve of the North American Caddisfly Genera (Trichoptera). 2nd Edition. University of Toronto Press. Toronto. Ontario. 457p.

Wood, D.M. 1963. The Black Flies (Diptera: Simuliidae) of Ontario. Part II Larval Identification with descriptions and illustrations. Proc. Entomol. Soc. Ontario.

Wright, J.F., Moss, D., Armitage, P.D., and M.T. Furse. 1984. A preliminary classification of running-water sites in Great Britain based on macroinvertebrate species and the prediction of community type using environmental data. Freshwat.Biol., 14: 221-256

Wrona, F.J., Culp, J.M., and R.W. Davies. 1982. Macroinvertebrate subsampling: a simplified apparatus and approach. Can.J.Fish.Aquat.Sci., 39: 1051-1054.

Yan, N. D., W. Keller, K. M. Somers, T. W. Pawson, and R. E. Girard. 1996. Recovery of crustacean zooplankton communities from acid and metal contamination: comparing manipulated and reference lakes. *Can. J. Fish. Aquat. Sci.* 53:1301-1327.

Zar, J.H. 1984. Biostatistical Analysis, Second Edition. Prentice Hall, Englewood, NJ.

7 Measurement of Supporting Environmental Variables

7.1 Introduction

A number of key environmental variables must be measured to aid in the interpretation of both the benthic and fish survey data. In addition, there are site-specific variables which may be measured where applicable (Tables 7.1a,b, and c). If there is a specific concern (e.g. potential eutrophication and nutrients), the approach would be to determine if the variable would be feasible to measure based on the effluent concentration and the known or predicated dilution factors in the receiving environment. Cost savings can be realized by approaching the analysis so that those samples with the likely highest concentrations are measured first, with measurements stopping when a pattern of undetectable samples is analyzed.

7.2 Receiving Water Monitoring

7.2.1 Sample Collection

The following procedures are recommended for collecting water samples. Sample bottle types are outlined in Table 7.2.

- For collection of surface water samples, samples should be collected directly into sample bottles, at a depth of about 15 cm, with the bottle mouth facing up-current away from the sampler's hand. Prior to collection of the sample, bottles should be opened under water, and should be triple rinsed with the sample water. Persons collecting samples should wear unlined latex or nitryl gloves to avoid contamination of the sample.

For depth samples, a peristaltic sampler should be used in preference to other types. If other types are used, they should be teflon-coated. Sampling should proceed from the least contaminated to the most contaminated station with two thorough rinses of pesticide-grade methanol and distilled water between stations. Laboratory blanks of the samplers should be run before and after use to demonstrate that no contamination is imparted to samples.

Table 7.1a Supporting environmental variables in water for freshwater (F) and marine (M) habitats

Key Variables	Site-specific Variables	Habitat (F,M)	Justification
Dissolved Oxygen		F, M	Dissolved oxygen can be decreased by mill effluents due to BOD.
Water Temperature		F, M	Mill effluents may cause increases in water temperature in receiving waters.
Salinity		M	Changes in estuarine conditions may affect benthic communities.
pH, Conductivity, hardness	Alkalinity	F	Provides information for water quality.
	Nitrate-Nitrite, Ammonia, and Total Nitrogen	F	Nitrogen is often a secondary limiting nutrient in freshwater.
	Sodium	F	Sodium may be a good tracer in freshwater.
	Dissolved Organic and Particulate Carbon.	F	Carbon is a nutrient source for microbes. Mills discharge quantities of these carbon sources.
	Colour or Turbidity	F, M	Mills may discharge effluents which are coloured. This may reduce the light available for primary production.
	Soluble Reactive or Total Dissolved Phosphorus and Total Phosphorus	F	Phosphorus is often the limiting nutrient in freshwater. Mills may discharge P which could lead to nutrient enrichment.

Table 7.1b Supporting environmental variables in sediment for freshwater (F) and marine (M) habitats

Key Variables	**Habitat (F,M)**	**Justification**
Sediment Carbon/Nitrogen ratio	M	C/N ratio is affected by terrestrial (wood) versus aquatic plant sources of marine deposition
Sediment Total organic carbon or Loss on ignition	F, M	Effects may be related to inputs from organic material.
Bed structure - Particle size analysis: Matrix: framework ratio, degree of embeddedness, texture, colour and thickness of layers in cores	F, M	Differences in physical structure of the habitat can influence invertebrate community structure.
Sediment Eh (redox)	M	Sediment redox provides an indication of the anoxic/oxic boundary in sediments.
Sediment total sulfides	M	Sulfides in marine sediments indicate the extent and nature of microbial response to organic enrichment.

Table 7.1c Other supporting environmental variables

Key Variables	**Site-specific Variables**	**Habitat (F,M)**	**Justification**
Latitude, longitude		F, M	Provides station location information.
Water depth		F, M	Water depth has a major effect on invertebrate communities.
	Algal biomass as chlorophyll or ash-free dry mass	F	Algae in the water column (lakes) or on benthic substrates provides a food source for higher trophic levels.
	Current velocity	F	Can provide information on equivalence of sampling stations.
	Bankfull channel width	F	Provides habitat structure information
	Other chemical tracers of mill effluent (in water, fish or sediment)	F, M	Provide evidence of effluent exposure.

Table 7.2 Handling recommendations for effluent and receiving water samples

Analysis	**Container**	**Volume**[1]	**Preservative**	**Maximum Storage**
TOC	HDLPE[2]	100 mL	H_2SO_4 to pH 2	10 days
DOC	HDLPE	100 mL	Filter, then H_2SO_4 to pH 2	10 days
Chlorinated Phenols, Guaiacols, Catechols	Amber Glass	1 L	4°C Filter, then H_2SO_4 to pH 2 for chlorinated phenols only	7 days to extraction
Resin and Fatty Acids (chlorinated and nonchlorinated)	Amber Glass	1 L	4°C or 0.5 g ascorbic acid plus 2 pellets NaOH	7 days or 14 days
Dioxins	Amber Glass	4 L	4°C	30 days
pH	HDLPE	1 L	4°C	Immediately
TSS, BOD_5	HDLPE	1 L	4°C	4 days
Nutrients (Total P, NO_2+NO_3 NH_3-N, TKN)	Glass	250 mL	H_2SO_4 to pH 2; 4°C for total P only	10 days
Chlorate	HDLPE	100 mL	4°C	28 days
Sublethal toxicity	Non-toxic Material[3]	2 L	Exclude air	3 days

[1] Suggested sample size
[2] High Density Linear Polyethylene
[3] Polyethylene or polypropylene is recommended.

Refer to NLET 1994a,b;1996

7.2.2 Sample Storage and Shipping

Table 7.2 summarizes recommended sample handling procedures and containers for the different analytical variables that may be included in the EEM monitoring program, e.g. bottle types for each variable or group of variables. These materials are considered as the best materials for the specific variable groups. Where appropriate, preservatives should be added to the sample bottle immediately on completion of the 24-hour composite sample collection. The actual sample volumes required may vary depending on the needs of the laboratory.

It is recommended that samples be cooled to 4°C during collection and stored at the same temperature for shipping to minimize degradation. Samplers should also be refrigerated, and shipping coolers equipped with ice packs or bagged ice to ensure that samples are kept cold. Speed in shipping (within 24 - 48 hours maximum) by pre-arranged transport is recommended.

7.2.3 Toxicity

Receiving water collected for toxicity tests should be pumped from the appropriate sampling depth using clean polyethylene tubing into a clean polyethylene or polypropylene container (typically a new food-grade polyethylene bag liner inside a 20-L bucket).

7.3 Sediment

7.3.1 Freshwater Environments

Dredges are commonly used to collect sediments (Klemm *et al.* 1990; Tetra Tech, Inc. 1987). For dredges and grabs, the following procedures are recommended to collect a representative sample for chemical analysis:

- A 5-cm layer of surficial sediments in freshwater systems (2-cm in marine systems) is recovered using a methanol-rinsed stainless steel spoon for chlorinated organics or plastic/teflon for metals.
- A colour photograph of each grab or core aids interpretation.

For cores, as described in Table 7.3, the same procedures are used except that the exterior "smear" zone of the core should be removed to avoid contamination. Several cores or dredges may be required to obtain sufficient samples to fill the sample containers.

Chemistry samples should be collected into 500-mL amber glass wide-mouthed sample jars (see Table 7.2 for preparation) with Teflon-lined lids. Similar containers can be used for toxicity testing samples, but polytetrafluoroethylene (PTF) is also acceptable.

Chemistry sample containers should be three-quarters full and frozen as soon as possible after collection until analyzed. Toxicity test samples should be refrigerated at 4°C (under anoxic conditions, if appropriate) until analyzed, but not frozen. Toxicity testing should commence within two weeks if possible.

Particle size determination is important for several reasons. It may tell us about the origin of sedimentary materials, and, in particular, about the dynamic conditions of transport and deposition. Also, in conjunction with chemical and biological analyses, particle size distribution help reflect or explain these other types of analyses.

In terms of particle size characterization, the rating scale of choice for most environmental studies remains the Wentworth Classification:

Clay: <0.0039 mm
Silt: 0.0039 mm-0.062 mm
Sand: 0.062 mm - 2.0 mm
Gravel: 2.0 mm - 16.0 mm

This rating scheme will be used in EEM studies.

7.3.2 Marine Environments

Sediment chemistry and particle size can be collected from the top of undisturbed grabs used for benthos collection. They can also be measured from samples collected by coring (see Holme and McIntyre 1984 for suitable coring devices). Core samples best preserve the surface sedimentary layer and are best taken by a SCUBA diver holding a core liner tube. The tube is pushed into the sediment as far as is possible, then capped at both ends

Table 7.3 Characteristics of coring devices (after Klemm *et al.* 1990)

KB Core Samplers

Habitats and substrates sampled: freshwater rivers, lakes, estuaries; soft sediments only, 40% silty clay.

Effectiveness: permits analysis of stratification in quantitative and qualitative samples; uses 5.08 cm (2 inch) pipe core tube; used in shallow to medium shallow water up to 30.5 m (100 feet) or deeper.

Advantages: samples a variety of substrates up to harder types; sampling tube can be modified for various diameters up to 100 cm^2 substrate surface; least disturbance to water/bottom interface; standard and heavy models available; wide variety of core tubes, liner tubes, core catchers, and nosepieces.

Limitations: gravity operated; samples limited surface area; standard KB core sampler head, without core tube weights approximately 8 kg (18 pounds), but additional weight can be added to sampler; requires boat and powered winch.

Ballchek Single and Multiple Tube Core Samplers

Habitats and substrates sampled: same as KB core sampler.

Effectiveness: samples deep burrowing organisms in soft sediment, particularly effective for sampling oligochaetes; uses 5.08 cm (2 inch) or 7.62 cm (3 inch) pipe core tube; used in shallow or deep waters, 3 m to 183 m (10 to 600 feet); multiple core sampler weight approximately 38 kg (84 pounds); check valves work automatically, preventing loss of sample.

Advantages: good penetration in soft sediments; small volume of sample allows for greater number of replicates to be analyzed in a short period of time; single or multiple (four) core tube sampler available; 7.62-cm (3-inch) pipe for larger cores and/or deep water lakes and oceans available; wide variety of core tubes, liner tubes, core catchers, and nosepieces.

Limitations: heavy device, approximately 38 kg, requires boat and winch; gravity operated; does not retain sand unless bronze core retainers are used, which require additional weight to insure penetration; may disturb water-sediment interface due to pressure wave.

Phelger Core Samplers

Habitats and substrates sampled: same as above core samplers.

Effectiveness: similar to KB core samplers.

Advantages: similar to KB core samplers.

Limitations: gravity operated or can be messenger operated with a suspension-release device; styles and weights vary among manufacturers, some use interchangeable weights, between 7 to 35 kg, others use fixed weights up to 41 kg; length core taken varies with substrate texture.

Box Core Samplers

Habitats and substrates sampled: same as above core samplers, also oceans.

Effectiveness: same as above core samplers; samples a surface area of 100 cm^2 and a sediment depth of 20 cm.

Advantages: same as above core samplers.

Limitations: same as above core samplers; also deployed from ships or other platforms; diver-collected cores are preferred.

Hand-Operated Core Samplers

Habitats and substrates sampled: same as above core samplers.

Effectiveness: sampled by hand or by diver.

Advantages: can be used in shallow water. In deep water, can be used with a diver, usually a trained biologist, who can collect and recognize substrate and bottom changes to stratify sampling; can be used with extension handles of 5, 10 or 15 feet; used with pipe fitting for driving from a pontoon boat, dock or bridge.

Limitations: limited area sampled.

before withdrawing it and returning it upright to the surface. Alternatively, the sample can be taken with a free-falling corer deployed from an anchored vessel (e.g. a Kajak type) if the sediments are sufficiently soft. After the metal part of the corer is forced into the sediment by the weight of its free fall the open end of the device is closed by spreading of an "egg shell" catcher and top opening of the core tube by a messenger operated rubber plunger. Core liner tubes of plastic, 5 x 50cm, are suitable and must be prepared for use (as described in Wildish *et al.* 1990). This involves drilling the tube in a spiral pattern so that holes occur at 5-cm intervals down the tube. The size of the holes is chosen so that a combination redox electrode and a cut-off 10 mL syringe can be pushed through the hole to obtain a sample for sulphide determination. During sampling the holes should be covered with duct tape so that they do not leak.

It is considered essential to use only those cores for further processing which have retained an intact sediment-water interface, without any form of leaking in carrying the core to the surface.

7.3.3 Critical Effects for Sediment Chemistry in Marine/Estuarine Environments

In marine systems, pulp mill effluent can affect benthos either by toxicity from chemical factors, or by organic enrichment. Determining critical effects of either type of effect are difficult. One of the most recent and comprehensive reports outlines recommended and tested assessment levels for reference biotic and physical factors related to environmental contaminants in Puget Sound, Washington (Striplin Environmental Consultants, Inc. 1996). Sediment quality standards for a series of sediment metals and organic contaminants have also been used nationally and in British Columbia (British Columbia Ministry of Environment 1991, CCME 1995, Nagpal and Pommen 1993, Tetra Tech., Inc. 1986). These guidelines may provide useful standards for comparison with results from coastal pulp mill benthos surveys.

Work done on organic enrichment from other industries, notably fish-farming, have shown what basic sediment chemistry and physical factors are of most importance to track and understand organic enrichments effects on biota. In general, the degree of impact on biota is related to the loading of usable organic carbon, related oxygen demand of sediments (chemical and biological) and the rate at which oxygen can be delivered or renewed in sediments. Therefore, Hargrave *et al.* (1993, 1997) have examined a series of sediment factors designed to determine these various effects, and have concluded for fish farms that those which are most indicative of enrichment and biotic effects include:

- sediment Eh (redox)
- total sulfides (S=), with important trends in, or relationships to TOC

- sediment particle size
- BOD
- CO_2 release
- biomass of deposit feeders

Although the character of the carbon and nitrogen from pulp mills will be different from that at fish farms, many of these factors will be equally as valuable to measure for marine and estuarine pulp mills in Canada. Therefore, it is suggested to use C/N ratios to indicate the source and type of terrestiral versus marine organic material in the sediments, so that this can be related to biotic factors.

Sediment TOC and particle sizes were measured consistently for soft, subtidal substrates in Cycle 1. TOC levels near diffusers ranged up to 6-7% in cycle 1, whereas normal marine sediments would be expected to be closer to a maximum of 4% for very fine sediments (see Washington State criteria for sediment quality standards applied to fish farms). However, natural levels of 6% were measured at one remote (unaffected by other human inputs) reference area on the west coast of Vancouver Island. Although it is not currently possible to determine important TOC levels which are related to the other chemistry and biotic factors at different impact levels at pulp mills, this type of measure has been used in combination with sediment particle size as a performance criteria for fish farms in Washington State (see Table 7.4). In addition, there is a clear relationship between redox level and TOC levels above 4% in B.C. fish farms (Burd 1997). TOC levels also declined notably along a gradient away from most of the fish farms, suggesting it is an important factor to be included.

Table 7.4 TOC sediment criteria for determining "normal" sediment conditions for monitoring in proximity to fish farms in Washington State

% silt/clay	TOC (%)
0-20	0.5
20-50	1.7
50-80	3.2
80-100	2.6

The ratio of C/N has also not been related to impacts on biota at pulp mills, but is an important chemical measure of the degree and type of nutrients from terrestrial wood in sediments, versus more natural marine or estuarine sources. Cellulose from trees has a much higher relative carbon content than other organic sources. At present, the only factors which can be used for criteria for sediment quality include Eh (redox) and S=

(total sediment sulfides). Redox indicates the oxygen condition of sediments, while sulfides indicate the organic breakdown of material occurring.

Therefore, a grading of impacts may be designed in the following way (Poole *et al.* 1978; Hargrave *et al.* 1995):

1) Gross impact

Impact is present and unacceptable in areas where natural anoxia does not occur due to poor basin flushing, when high TOC levels are combined with anoxic surface sediments (Eh <0, S= >2000), high C/N, no macrofauna or bioturbation. This would trigger a change in effluent standards IF: a sediment profile shows the chemical factors are not just the result of historical impacts from the mill prior to recent improvements in effluent standards. For instance, if there is no accumulation of natural sediments over historical fibre mats, chances are the impact is an on-going one. However, if the surface layer is obviously an improvement over conditions at depth in sediments, then a historical impact can be inferred, which may not trigger a change in effluent standards, but may require remediation of the site.

2) Moderate-high impact

Moderately high TOC and C/N in sediments which are not naturally anoxic, Eh about 0-100, S= >200, total abundance of macrofauna < 50% stations further away from diffuser, taxa number 10 or less; biomass reduced by >50% from stations further away. This combination can be important because there are areas of natural anoxia (such as Alberni Inlet in B.C.) where the organic content and sulfides are not high. Any biotic reductions in this area would be the result of natural factors rather than the mill.

3) Low impact or enrichment

TOC and C/N relatively normal for the sediment type and natural oxygen conditions, Eh=100-200, S= 50-200, total abundance as high or higher than stations further away from diffuser, taxa number >50% that of stations further away, biomass >50% as high or even higher than stations further away.

4) **Normal**

TOC and C/N relatively normal for the sediment type and natural oxygen conditions, Eh>200, S=<50, total abundance, biomass and taxa number greater than stations further away from diffuser. Note that for naturally anoxic conditions these standards do not apply.

7.4 Quality Assurance in Analyses

Analytical quality assurance defines the way in which tasks are to be performed to ensure that data meet pre-defined data quality goals. These tasks include not only the analysis itself but all aspects of sample handling and data management. The following QA requirements as outlined by CAEAL (1991) are recommended by analytical laboratories undertaking chemical analyses as part of the EEM program.

7.4.1 Sample Handling and Tracking

a) **Chain of Custody** Chain-of-custody forms should be used in the transportation of samples, especially in cases where several contracted parties are involved in the sampling, shipping, and analysis of the samples. An example of a chain-of-custody form is presented in Figure 7.1.

b) **Sample Inspection** The condition of each sample should be noted upon receipt. Discrepancies between required sample conditions and the observed conditions should be recorded in a logbook or on a computer file. The samples should be preserved immediately if submitted unpreserved, and a record made of the preservation methodology.

c) **Sample Tracking** Samples should be assigned a unique number or code to identify the sample in a tracking system. The sample tracking system should identify the sample, the source, the date of receipt, analyses, due date, and any other pertinent information. A computerized laboratory information management system (LIMS) is recommended for tracking samples in laboratories processing large numbers of samples for a variety of clients.

d) **Sample Storage** Samples should be stored in an assigned location in a refrigerator or sample storage area accessible only to authorized personnel. Samples should be refrigerated at 4°C, where applicable, and removed only for inspection, logging, and analysis. The temperature of the refrigerator should be measured and recorded daily.

Figure 7.1 Example of a chain-of-custody record (modified from U.S. EPA 1988)

Survey				Samplers (signature)					
Station Number	Station Location	Date	Time	Sample Type Water		Preservative added (initial)	Seq. No.	No. of Containers	Analysis Required
				Comp	Grab				

Relinquished by (signature)		Received by (signature)	Date/time
Relinquished by (signature)		Received by (signature)	Date/time
Relinquished by (signature)		Received by (signature)	Date/time
Relinquished by (signature)		Received by (signature)	Date/time
Dispatched by (signature)	Date/time	Received for laboratory by:	Date/time
Method of shipment:			
Distribution Orig. - accompany shipment 1 copy - survey coordinator field files			

7.4.2 Good Laboratory Practices

Well established, good laboratory practices (GLPs) should be followed. The following is a brief listing of recommended laboratory practices. A description of GLPs can be found in greater detail in ELAP (1988).

- Records on reagent preparation should be maintained in a log book. Prepared reagent containers should be labelled with the reagent, its date of preparation, expiry date, and the person responsible.
- Instruments should be maintained or serviced on a regular basis. Maintenance records should be kept in a log book.
- Written instructions should be available for all instruments.
- Standard procedures for cleaning glassware and containers should be followed.
- Routine checks of the purity of the distilled water should be conducted and documented. Distilled/deionized water should be checked on a conductivity meter at least daily.
- Chemical reagents should meet the purity requirements of each analytical method.
- Reagents and solvents should be stored according to the manufacturer's directions.
- Working standards and stock solutions should be checked to determine changes in concentration.
- Reagents should be prepared and standardized against primary reference standards.
- The temperatures of all refrigerators and incubators should be checked daily and temperature excursions should be recorded.
- Each oven should have a dedicated thermometer and the temperature should be checked prior to and following each use.
- Proper volumetric glassware should be used.
- Glassware should be cleaned according to specifications of the method.
- Gas cylinders should be replaced at 700-1400 kPa.
- Laboratory personnel should have appropriate training in analytical laboratory procedures, and in the particular analysis for which they are responsible.

7.4.3 GC/MS (Gas Chromatography/Mass Spectroscopy) Tuning and Performance Checks

For analyses using mass spectrometry detection, instruments should be tuned to provide a mass calibration of the instrument (CAEAL 1991; U.S. EPA 1988). Instrument tuning is performed upon method set-up or as a corrective action. CAEAL (1991) recommends use of PFTBA (perfluorotributylamine, FC43) for tuning and has set specific performance criteria. Perfluorokerosene (PFK) is normally used to calibrate high resolution mass spectrometers. In addition, GC/MS system performance checks using specific compounds should meet the performance requirements presented in CAEAL (1991).

7.4.4 Organic Analytes Identification

Method principles for a variety of organic analytes by various analytical methods and for different sample types are presented in Tables 7.5 and 7.6. Acceptance criteria for GC/MS analyte identification and additional procedures for identification of analytes by gas chromatography and HPLC (High Pressure Liquid Chromatography) are outlined in CAEAL (1991).

7.4.5 Instrument Response Checks

a) **Inorganic Analyses** Instrument response (absorption, counts, peak height, etc.) can be checked before calibration or analysis proceeds by comparing the response of a check sample used on an ongoing basis, to historical data. Instrument response can be charted in control charts with a warning limit set at ± 2 standard deviation and an out-of-control limit of ± 3 standard deviations.

b) **Organic Analyses** Instrument response checks can be conducted with organic analyses by daily comparison of the internal standard response or an analyte response of a mid-range standard to historical data. Typically, an analyte or analytes that are most sensitive to instrument response are chosen for this check. Warning and control limits are set at ±2 standard deviations and ±3 standard deviations (SD), respectively. Other standard external checks can be used depending on the type of instrumentation.

7.4.6 Instrument Calibration Requirements

a) **Inorganic Analyses** Calibration is recommended for all inorganic methods with a minimum of three concentration levels (5 recommended) of standards, plus a blank, so that all samples fall within the calibrated range (CAEAL 1991). In cases where instrument software limits the calibration to two points, the calibration range should not exceed the linear range of the instrument. Unless it is demonstrated that no bias is introduced, standards must be matched to the matrix of the samples.

b) **Organic Analyses** Calibration for analyses using GC/MS and GC is achieved with five concentration levels of standards to cover an expected linear range generally by using internal standards and by calculating an average response factor for each

Table 7.5 Summary of analytical method principles for effluents and receiving waters

Parameter	Sample Preparation	Instrumental Method	Matrix Detection Limits		Precision[1]	Accuracy[2]	References
			Effluent	Water			
Chlorate	Filter or centrifuge sample to remove solids which may plug ion chromatograph.	Ion chromatography using Dionex AS9 column or equivalent.	0.5 mg•L^{-1}	20 µg•L^{-1}	± 10%	80-120%	U.S. EPA (1991)
Chlorinated Dioxins and Furans	Spike samples with isotopically labelled PCDD/F surrogates; extract with solvent and concentrate; clean-up as required including back extraction with acid/base, gel permeation, alumina, silica gel, activated carbon; concentrate to known micro-volume and add isotopically labelled performance standards immediately prior to instrumental analysis.	High resolution gas chromatograph coupled to high resolution mass spectrometer with computerized data system (HRGC/HRMS/DS); minimum two characteristic ions per congener group; 5-point calibration over linear range of MS	TCDD/F P5CDD/F H6CDD/F H7CDD/F OCDD/F (Only In Effluent)	2-4 pg•L^{-1} 4-8 pg•L^{-1} 4-8 pg•L^{-1} 6-12 pg•L^{-1} 8-16 pg•L^{-1}	±20%	Native Compound Spike 80-120% Surrogates 40-130%[3]	Environment Canada (1992b) U.S EPA (1990a)
Chlorophenols, Chloroguaiacols, Chlorocatechols	Spike samples with appropriate surrogates; in-situ acetylation followed by solvent extraction OR acidification to pH 2, solvent extraction/ concentration, extract acetylation or methylation; addition of internal standard immediately prior to instrumental analysis.	HRGC/LRMS/DS in Selected Ion Monitoring (SIM) mode using a minimum of two ions (1 quantitation, 1 confirmation) in correct abundance ratio within established retention time windows.	0.2-1 µg•L^{-1}	0.02-0.05 µg•L^{-1}	± 20 %	70-110%[3] (Phenols and Guaiacols) 40-90%[3] (Catechols)	NCASI (1986) Carron and Afghan (1989) Lee *et al.* (1989) U.S. EPA (1991) Alberta Environment (1991) Morales *et al.* (1992)
Nutrients-Ammonia	For manual determination, effluent samples and some waters will require distillation.	Manual measurement of indophenol blue at 630 nm OR Automated distillation, reagent addition and colour measurement at 630 nm on segmented flow analyzer.	0.05 mg•L^{-1}	0.005 mg•L^{-1}	± 10%	95-105%	NLET (1994a) APHA (1989, 1992, 1995)
Nutrients - Nitrate Plus Nitrite	None	Manual or automated reduction cadmium, addition of reagents to form azo dye, measurement of colour at 520 nm.	0.1 mg•L^{-1}	0.01 mg•L^{-1}	± 20% at 0.5 mg•L^{-1}	95-105%	NLET (1994a)) APHA (1989, 1992, 1995)
Nutrients - Total Kjeldhal Nitrogen	For manual determination, digestion with sulphuric acid/potassium sulphate/ mercuric oxide solution required prior to distillation and instrumental determination of ammonia.	Manual measurement of indophenol blue at 630 nm OR Automated digestion, distillation, reagent addition, and colour measurement at 630 nm by segmented flow analyzer.	0.02 mg•L^{-1}	0.02 mg•L^{-1}	± 10%	95-105%	APHA (1989, 1992, 1995)

1. Acceptable performance for duplicates based on 1 standard deviation derived from historical data.
2. Acceptable performance for recovery of reference material or spikes, based on 1 standard deviation derived from historical data.
3. Control limit (± 3 SD) based on recovery of surrogate compound spikes.

Table 7.5 Summary of analytical method principles for effluents and receiving waters (continued)

Parameter	Sample Preparation	Instrumental Method	Matrix Detection Limits		Precision[1]	Accuracy[2]	References
			Effluent	Water			
Nutrients - Total P	For manual determination, digestion by sulphuric acid/potassium persulphate OR sulphuric acid/potassium sulphate/ mercuric oxide if TKN being analyzed at same time	Manual measurement of molybdate blue complex at 880 nm OR Automated digestion, reagent addition, and colour measurement at 380 nm by segmented flow analyzer.	0.002 mg•L^{-1}	0.002 mg•L^{-1}	± 10%	95-105%	NLET (1994a) APHA (1989, 1995)
pH	None	Electrometric	NA	NA	± 0.05	± 0.1	NLET (1994a) APHA (1995)
Resin and Fatty Acids	Spike all samples with o-methyl podocarpic acid surrogate; adjust pH to 9 and extract with methyl t-butyl ether and concentrate; derivatize by methylation with diazomethane using tricosanoic acid addition as recovery check.	HRGC/LRMS in SIM Mode using minimum of two characteristic ions per compound in correct abundance ratios within established retention time windows.	2-5 µg•L^{-1}	0.2-0.5 µg•L^{-1}	± 20%	60-120%[3]	Alberta Environment (1990, 1991) Voss and Rapsomatiotis (1985) NCASI (1986) OMOE (1989)
BOD$_5$	Treat to remove interferences from residual chlorine, sulphide and adjust pH to 6.5-8.3; dilute samples to provide minimum D.O. depletion of 2.0 mg•L^{-1} and minimum D.O. concentration of 1.0 mg•L^{-1}; for pulp and paper mill samples use acclimatized seed.	Measure dissolved oxygen by stirring BOD probe calibrated against Winkler, or use Winkler.	5 mg•L^{-1}	2 mg•L^{-1}	± 15% (Inter-laboratory)	80-120%	NLET (1994a) APHA (1989)
Total Organic Carbon (TOC)	Acidify to pH 2 and purge with nitrogen to remove inorganic carbon.	High temperature combustion with infrared detection of CO$_2$ OR Ultraviolet/persulphate digestion with colourimetric or infrared detection of CO$_2$.	1.0 mg•L^{-1}	1.0 mg•L^{-1}	± 5	95-105%	NLET (1994a) APHA (1989, 1992, 1995)
Total Suspended Solids (TSS)	**Effluents and Fresh Waters:** Filter sample through washed, dried pre-weighed Whatman GFC glass fibre filter, and dry filter and solids at 105°C to constant weight	Determine weight change in filter gravimetrically.	2 mg•L^{-1}	2 mg•L^{-1}	± 10%	90-110%	APHA (1989) NLET (1994a)
	Marine Receiving Waters: Filter through pre-weighed Nuclepore 0.4 mm membrane filter, wash thoroughly with distilled water, dry at 80°C and re-weigh		NA	50 µg•L^{-1}	±20%	NA	Sundby (1974)

Table 7.5 Summary of analytical method principles for effluents and receiving waters (continued)

Parameter	Sample Preparation	Instrumental Method	Matrix Detection Limits		Precision[1]	Accuracy[2]	References
			Effluent	Water			
Algal Biomass		chlorophyll or ash-free dry mass					APHA (1995)
Alkalinity							APHA (1995)
Colour		platinum-cobalt method					APHA (1992, 1995)
Conductivity		Hydrolab, YSI (600-series) or CTD multi-probes					
Current Velocity		Global Flow Probe current meter					
Hardness							APHA (1995)
Dissolved Oxygen		Hydrolab, YSI (600-series) or CTD multi-probes. May need to use Winkler titration method for QA comparison to meters					
Salinity		In situ YSI meter (600-series) or salinity meter which indirectly utilizes conductivity and temperature					APHA (1992)
Dissolved Sodium		Flame atomic emission spectrophotometry					APHA (1992, 1995)
Temperature		Hydrolab, YSI (600-series) or CTD multi-probes.					
Turbidity		turbidimeter					APHA (1992, 1995)

Table 7.6 Summary of analytical method principles for sediment

Parameter	Sample Preparation	Instrumental Method	Matrix Detection Limits	Precision (% rel std dev)	Accuracy (% recovery)	Reference Methods
Chlorinated Dioxins and Furans	Air dry, weigh and spike with isotopically labelled surrogate standards. Extract prepared sample, concentrate and clean-up per effluent method.	HRGC/HRMS/DS per Effluent Method	TCDD/F 1pg•g^{-1} P5CDD/F 5pg•g^{-1} H6CDD/F 5pg•g^{-1} H7CDD/F 5pg•g^{-1} 0.1-0.5 mg•g^{-1} Depending on Isomer OCDD/F 10 pg•g^{-1}	± 20%	40-120%	U.S. EPA (1990a) U.S. EPA (1990b) Environment Canada (1992a) See Table 5.4 Chlorophenol method
Chlorophenol, Chloroguaiacol, Chlorocatechol Isomers	Weigh out wet sediment and spike with surrogate; extract and concentrate; derivatize with large excess of acetic anhydride.	HRGC/LRMS per Effluent Method.	10-50 ng•g^{-1} depending on isomer	± 20%	70-110% 50-90%	Birkholz *et al* (1988) Hynning *et al.* (1989) Lee *et al.* (1989)
Chlorinated Resin and Fatty Acids	Extraction as per Chlorophenols; no clean-up; derivatize as per Resin and Fatty acids in Effluents.	HRGC/LRMS per Effluent Method	0.1 μg•g^{-1}	± 20%	70-120%	See Table 5.4 Morales *et al* (1992)
Particle Size	Freshwater: Dry and sieve standard series. Marine: Wash through 63 mm sieve; <63 mm fraction recovered for weighing; >63 mm fraction dried and sieved in standard series.	Determine percent of each sieve fraction gravimetrically. OR laser particle size method using Malvern 2600L laser particle size analyzer.	±0.1%	NA	NA	Walton (1978) Tetra Tech (1986)
Total Sulphide	5 ml of sediment is squeezed into airtight container with 5ml of sulphide antioxidant buffer (SAOB) solution.	MV readings are made within 3 hours, using silver sulphide and double junction reference electrodes				Tetra Tech (1986)
Volatile Residue (SVR) or Loss on Ignition (LOI)	Sample is oven-dried	Gravimetric analysis after ignition in muffle furnace at 550°C				APHA (1992) Tetra Tech (1986)

Table 7.6 Summary of analytical method principles for sediment (continued)

Parameter	Sample Preparation	Instrumental Method	Matrix Detection Limits	Precision (% rel std dev)	Accuracy (% recovery)	Reference Methods
C:N (organic C to total N)		Measure TOC as per below, measure total N as per total C				Calvert *et al.* (1995) MacDonald *et al.* (1991)
Eh	Measured with sediment still in core tube.	Calibrated probe is pushed into sediment.				Hargrave *et al.* (1995)
Total Organic Carbon Total Organic Nitrogen	Sample is dried, ground to pass #80 sieve; treat with sulphurous acid to remove inorganic carbon and dry;	(e.g. Perkin-Elmer CHN analyzer) OR Determine COD and divide result by 2.6. OR use USEPA (1986) Method 9060A - total carbon analyzed using Leco induction furnace, carbonate carbon analyzed using a Leco gasometer. TOC reported as difference between total C and carbonate C	NA	NA	NA	American Soils Association (1982) Hargrave *et al.* 1997 U.S. EPA (1986)

analyte over five concentration levels. These average response factors are then used to calculate the unknown concentrations in the samples. A calibration is considered valid if the relative percent standard deviation of a multiple calibration is less than or equal to 30% for all variables. The relative percent standard deviation is defined as:

$$\% \text{RSD}_A = \frac{\text{SD}}{\text{Avg RRF -A}} \times 100$$

where: $\% \text{RSD}_A$ = relative percent standard deviation of the response factor for variable A

SD = standard deviation of the response factor of variable A over the five concentration levels, and

Avg RRF_A = average response factor for variable A over the five concentration levels.

If this criterion is not met, corrective action should be taken and the initial calibration should be repeated. Analysis of samples should not proceed until a valid calibration is established.

7.4.7 Analytical Documentation

Documentation of all aspects of the analysis is recommended to confirm the quality and reliability of the analytical results. Storage of sample results and data associated with the analyses in hard copy or on computer file (with back-up) is recommended for at least five years after analysis (CAEAL 1991). Information for each sample or batch of samples information is recommended on the following:

a) **Method Detection Limits (MDLs)** If MDLs are different from the laboratory-determined MDLs (due to interference, dilutions, etc.), this should be recorded.

b) **Sample Storage Times** Records should be kept on the sampling date, date of receipt, date of sample preparation, and date of analysis. This information is normally handled as part of the sample tracking process.

c) **Instrument Performance and Maintenance** A log should be kept of instrument performance including records of tuning and instrument response. Maintenance or service records should be kept for each instrument.

d) **Quality Control Samples** Records of duplicate analyses, blanks, spiked blank recoveries, surrogate recoveries, matrix spike recoveries and results from certified reference materials, and records of calibration and calibration checks should be maintained.

e) **Sample Reception, Preparation, and Analysis** All anomalies in delivery, storage, condition, preparation, and analysis of samples should be recorded. These include any deviations from standard operating procedures.

7.5 Analytical Quality Control

Analytical quality control procedures are designed to demonstrate statistical control over calibration, precision, accuracy/bias, and recovery (CAEAL) 1991.

Statistical control over these parameters, can be demonstrated by running specific quality control samples during each analytical run (see Tables 7.7and 7.8 or inorganic and organic analyses, respectively). The results of these QC samples are compared statistically with confidence intervals calculated from historical data. These confidence intervals or control limits are normally calculated at three standard deviations of the mean of the controlled variable. Warning limits are frequently set at two standard deviations. Indicators of a run considered out of control include the following:

- two successive results for method blanks, laboratory duplicates, standard reference materials, spiked blanks, calibration control samples, or organic surrogate recoveries
- one of these results outside the control limits.

QC data can be plotted on appropriate control charts. Control charts are graphic presentations of the QC data as a function of time or consecutive run number. Control charts demonstrate trends in time and provide graphic evidence of long-term statistical control of the analysis. Control limits and control charts are described in detail in ASTM (1986).

7.5.1 Calibration Control

Statistical control over calibration can be confirmed between runs by means of two control standards, A and B, and within-run by means of blanks and midrange standards.

a) **Between-run Calibration Control** Two control standards, A and B, can be used to analyze and control between-run changes in calibration, once at the beginning of each analytical run. These standards are made up and maintained independently of the calibration standards and are normally chosen to be about 80% and 20% of full-scale, respectively. Results are accumulated over many runs and the sums (A + B) and

Table 7.7 Recommended quality control for inorganic analyses

QC Sample	Frequency	Control Limits	Reference
Method Blank	1 per 10 samples	Less than Detection Limit or Less than 0.1 times sample levels	CAEAL (1991)
Laboratory Duplicates	1 per 10 samples	See Tables 6.1 and 6.2	CAEAL (1991)
Matrix Spikes	1 per 10 samples	75% to 125%	CAEAL (1991)
Spiked Method Blank	1 per batch or minimum 5%	75% to 125%	CAEAL (1991)
Standard Reference Materials	1 per batch or minimum 5%	See Tables 6.1 and 6.2	CAEAL (1991)
Calibration Control:			
• within run (blank and mid-range standard)	1 per 20 samples	10% drift max.	CAEAL (1991)
• between runs (20% and 80% full scale)	2 per run	± 5% of target value	King (1976)

Table 7.8 Recommended quality control for organic analyses

QC Sample	Frequency	Control Limits	References
Method Blank	1 per 10 samples	Less than Detection Limit or Less than 0.1 times sample levels	CAEAL (1991)
Laboratory Duplicates	1 per 10 samples	See Tables 6.1 and 6.2	CAEAL (1991)
Matrix Spikes (Authentic Compounds)	1 per 10 samples	90% of compounds meet limits in Table 4 of CAEAL (1991)	CAEAL (1991)
Surrogate Spikes (for GC and GC/MS)	100%	60% to 120%	CAEAL (1991)
Spiked Method Blank	1 per batch or 5% minimum	70% to 120%	CAEAL (1991)
Standard Reference Materials	1 per batch or 5% minimum	See Tables 6.1 and 6.2	CAEAL (1991)
Internal Standards	100%	NA	CAEAL (1991)
Calibration Control: • within run (mid-run standard)	1 per 12 hours or 20 samples whichever is more frequent	± 25%	CAEAL (1991)

differences (A - B) are plotted on control charts. During a specific run, a significant change in the sum (A + B) from the historical mean implies that a significant change in intercept has occurred, other factors remaining constant. A significant change in the difference (A - B) implies a significant change of slope, other factors remaining constant. Control and warning limits for A - B are calculated for the mean and the standard deviation of the population of differences:

$$\text{Upper and Lower Warning Limits (UWL, LWL)} = X_{A-B} \pm 2\ SD_{A-B}$$

$$\text{Upper and Lower Control Limits (UCL, LCL)} = X_{A-B} \pm 3\ SD_{A-B}$$

Control and Warning limits for A + B are similarly calculated using the same standard deviation:

$$\begin{aligned} \text{UWL / LWL} &= X_{A+B} \pm 2\ SD_{A-B} \\ \text{UCL / LCL} &= X_{A+B} \pm 3\ SD_{A-B} \end{aligned}$$

The run should not proceed until it is shown that A + B and A - B are within control limits. Control limits should not exceed ± 5% of the average value for A+B and A-B.

b) **Within-run Calibration Control (Inorganic Analyses)** Within-run changes in calibration attributable to slope and baseline drift should be checked at regular intervals. This can be accomplished by use of a mid-range standard and reagent blank run after every 20 samples. Control limits should be established by each laboratory for each procedure. The drift should not exceed 10%. If a greater drift is detected, the analysis should be stopped, the instrument re-calibrated, and samples run after the last acceptable check sample and blank are reanalyzed.

c) **Within-run Calibration Control (Organic Analyses)** In organic analyses by GC, within-run changes in calibration should be checked by injection of a mid-level check standard at a frequency of 5% or every 12 hours. This injection is compared to the initial calibration by calculating the percent deviation in the response factor of each analyte in the check standard to the average response factor determined during the initial calibration. If the relative percent difference is greater than 25%, the calibration check should be repeated. If the repeated check standard still has a relative percent deviation greater than 25%, corrective action is recommended.

7.5.2 Precision

Precision is the degree of variation among individual measurements of the same variable using a specific analytical method and is usually expressed as the standard deviation of replicates (U.S. EPA 1990c). Statistical control of analytical precision is maintained by analyzing within-run duplicates at a frequency of at least 10%. Laboratory duplicates are separate aliquots split in the laboratory from a single sample.

The absolute difference between within-run duplicates is compared to a control limit determined from historical data. To obtain these control limits, the results of duplicate analyses are accumulated over many runs and sorted according to concentration ranges.

Convenient concentration ranges are 0 to 20%, 20 to 50%, and 50 to 100% of full scale (King 1976). Within each concentration range, control limits for the absolute difference between within-run duplicates is determined from the formula:

$$UCL = D_4 \times R$$

where D_4 is a statistical factor and R is the mean difference between duplicates, 3.267 (ASTM 1986; Taylor 1987).

If the difference between laboratory duplicate analyses exceeds the upper control limit, the situation should be evaluated to determine the most appropriate corrective action.

7.5.3 Accuracy and Bias

Accuracy is the degree of agreement between an observed value and the true value as determined by analysis of an accepted reference material (U.S. EPA 1990c). The converse of accuracy is the degree of systematic error in the analysis or bias. Accuracy is controlled by means of method blanks and certified reference materials.

a) **Method Blanks** A method blank is an aliquot of reagent water equivalent in volume to the samples being processed and run in exactly the same manner as the samples. The method blank quantifies the level of contamination introduced to the samples during sample processing and analysis. Method blanks should be analyzed at a frequency of 10% or 1/run, charted, and controlled at ± 2 SD (warning limits) and ± 3 SD (control limits). If a method blank is judged out of control and contaminated, those samples processed with the blanks and greater than the detection limit should be repeated for the variable(s) affected. In general, a method blank is considered free of contamination if the analysis yields results less than the detection limit or less than 0.1 times the level found in all associated samples (CAEAL 1991).

b) **Standard Reference Materials** Standard reference materials (SRMs) are samples available in different matrices that have been extensively analyzed by several laboratories and have concentrations certified by standard-setting organizations such as the National Institute of Science and Technology, the U.S. EPA, the National Water Research Institute of Environment Canada and the National Research Council. When available, an SRM should be analyzed at a frequency of 5% or 1/run (Tables 7.8 and 7.9). The matrix and concentration of the SRM should be as close as possible to the samples being analyzed. The results of SRMs should be accumulated and control and warning limits determined as ±3 SD and ±2 SD, respectively.

7.5.4 Recovery

Recovery of the analyte over the entire analytical process is determined from matrix spikes, spiked blanks, and surrogate spikes.

a) **Matrix Spike** A matrix spike is a separate aliquot of a randomly chosen sample to which is added all the analytes of interest before processing of the sample. Analysis of a matrix spike gives an indication of the recovery efficiency obtained for the matrix particular to that sample. The sample should be spiked with all the analytes of interest at a concentration as close as possible to that concentration giving a response equal to the mid-level calibration standard. The spiking solution should be prepared from a stock source separate from that used for calibration. The recommended distribution of matrix spikes is 10% or 1/run (Tables 7.8 and 7.9). One method to calculate recovery is:

$$\%\text{ Recovery} = \frac{\text{Measured Conc. - Unspiked Conc.}}{\text{Spike Amount}} \times 100$$

The results of matrix spikes should be plotted on separate control charts for each matrix. In-house limits should be set on the basis of ± 3 SD on a minimum of 10 data points. In multi-parameter analyses, at least 90% of the analytes should have recoveries within the specified limits. Recoveries for inorganic analytes should fall within 75 to 125%. Recoveries for organic variables should fall within the limits specified in Table 4 of CAEAL (1991). If a matrix spike does not meet these criteria, the spike should be repeated. If the recoveries do not meet the criteria in the repeat analysis and there are no indications of other problems with the analysis, a matrix effect should be noted and reported.

b) **Spiked Method Blank** The spiked method blank is a separate aliquot of the same reagent water used for the method blank that is spiked with the compound of interest at a concentration as close as possible to the concentration of the mid-level calibration standard. The spiked method blank gives an indication of the reliability of a method without the matrix effects of real samples. The spiked method blank should be processed with and in the same manner as the samples. As with the matrix spike, the spiking solution should be prepared from stocks separate from those used for calibration. One method to calculate recovery is

$$\% \text{ Recovery} = \frac{\text{SMB - MB}}{\text{Spike Amount}} \times 100$$

where: SMB = Concentration measured in spiked method blank
MB = Concentration measured in the method blank

In-house recovery limits should be calculated for the spiked method blank based on ± 3 SD and a minimum of 10 data points (Tables 7.7 and 7.8). Recoveries for inorganic analyses should fall between 75% and 125%. Recoveries for organic variables should fall within 70% and 120%. If a spiked blank recovery does not meet the criteria established, the spike should be repeated. If the spike still does not recover, the samples related to the spike should be repeated. If insufficient sample remains for a repeat analysis, the results should be reported and flagged as suspect with an explanation.

c) **Internal Standards (Organic Analyses)** All analyses using GC should be performed using internal standards, or properly validated methods using external standards. An internal standard is a compound that behaves similarly in an analytical system as the compounds of interest, but is unlikely to be found in the sample. Internal standards are added at the same level to all samples, standards, and control samples prior to measurement but after sample preparation. All analyte responses should be normalized for the internal standard response to correct for instrument variability in response to such factors as varying injection volumes, temperature fluctuations, and final extract volume. The response of the internal standard in the sample measurement should be within 20% of the internal response of a calibration standard analyzed within the same 12-hour period (Table 7.8). If this criterion is not met, the sample should be repeated. If upon re-analysis the criterion is still not met, the sample results should not be corrected for internal standard response and should be flagged with an explanation.

d) **Surrogate Spikes (Organic Analyses)** A surrogate standard is a compound not expected to be found in the sample, that behaves similarly to the analytes of interest during sample preparation and analysis. Where applicable, surrogates should be added to all samples (including QC samples) before sample preparation to indicate method performance and sample matrix effects. Analyses run by GC/MS should have at least two surrogates while, those run by GC should have at least one surrogate. The amount of surrogate added to all samples should be the same as that added to the calibration solutions. In-house control limits for surrogate recoveries are based on ±3 SD on a minimum of 10 data points. In-house control limits for surrogate recoveries should be within 60% to 120%. If any surrogate is outside the expected recovery range, the sample should be reanalyzed. If, upon reanalysis, the surrogate recovery is still outside the permissible range, the results should be reported with a flag and an explanation.

7.5.5 Detection Limits

Detection limits should be reported as the method detection limit (MDL) as described by U.S. EPA (1984). The MDL is defined as the minimum quantity of an analyte that must be observed to justify the claim to have detected the analyte with a specified risk (normally 5% or 1%) of making a false detection.

One method to calculate the MDL is from the standard deviation of the analysis at the lowest concentration range:

$$\mathrm{MDL} = t_{0.05\,n-1} \times S \qquad (1)$$

where: $t_{0.05,\,n-1}$ is the one tailed value of Student's t for a 5% risk of false detection, n-1 degrees of freedom, and S is the standard deviation.

Ideally the standard deviation is calculated from low level replicate analysis on real samples having the same or similar sample matrix as the samples under consideration. This standard deviation can be calculated from a minimum of seven replicates in the same run using the standard statistical formula (U.S. EPA 1984). However, it is preferable to calculate S from between-within-run replicate pairs accumulated over many runs.

The standard deviation of low-level replicate pairs accumulated over a large number of analytical runs is:

$$S = \frac{\Sigma\, S\, D^2}{2n} \qquad (2)$$

where D is the individual replicate difference and n is the number of replicate pairs. A minimum of 40 replicate pairs is recommended (OMOE 1988). The value of either standard deviation is then entered in equation (1) to calculate the MDL.

7.6 Data Reporting Conventions

Established protocols for rounding-off analytical results should be followed. If too many figures are rounded-off before reporting, information is lost and real differences in the concentrations of samples from different locations or occasions may be concealed. Quality control procedures may be on a coarser basis than is desirable, or necessary, with the result that values of the mean, standard deviation, or other statistics of a set of results may be biased. Conversely, when too many significant figures are reported, relatively small, statistically insignificant differences may appear falsely large (Hunt and Wilson 1986).

The standard deviation of the analysis is the preferred criterion for deciding the number of significant figures (King 1989). The process of rounding-off should ensure retention of the digit that is in the same decimal position as the most significant digit in the calculated standard deviation. For example, if the analysis provides a value such as 12.345 and the calculated standard deviation based on within-run replicate analysis at this concentration level is 0.32, the result should be truncated to 12.3.

7.7 Method Principles - Physical and Chemical

Table 7.5 provides a summary of recommended method principles for some variables for effluents and receiving waters, and recommended performance criteria for precision and accuracy. Table 7.6 presents corresponding information for sediments and tissues. The criteria for precision are based on the relative percent standard deviation of the analysis and are derived from historical data on replicate analyses of samples in the range of 20 to 50% of full scale. Under most circumstances, the analyst should expect duplicates to differ by less than the percentage indicated.

For inorganic analyses, the criteria for accuracy are presented at ± one standard deviation of the percent recovery of a standard reference material based on historical data. Under most circumstances, a reference material should fall within the range indicated.

For organic analyses, the performance criteria are presented as actual control limits on spikes added to each sample (Table 7.8). These limits should be calculated at three standard deviations from the historical mean recovery of each spike in the same matrix as the sample. In the event that these limits are not met in a particular sample, the sample should be repeated. If the repeated sample still does not meet the recovery criteria, the results should be flagged to indicate matrix effects in the absence of any other indications of problems with the analysis.

Comments specific to selected methods are presented below.

7.7.1 Resin and Fatty Acids in Effluents

While interlaboratory validation for the extraction and derivatization method for resin and fatty acids (RFA) in effluents in Voss and Rapsomatiotis (1985) has not been completed, the method is recommended for EEM. However, GC/MS in SIM mode is recommended as the preferred approach for instrumental analysis to avoid interferences. An o-methyl podocarpic acid surrogate is recommended as a method recovery check, while tricosanoic acid should be added immediately prior to diazomethane derivatization as a methylation check. A heneicosanoic acid internal standard should also be used in instrumental analysis. Quantitation ions are those presented in NCASI (1986).

7.7.2 Chlorinated Dioxins and Furans in Sediments

Extraction of sediment is different from that for water, but the extract should be cleaned up and analyzed as described in the Environment Canada method for effluents (Environment Canada 1992a). For sediments, samples are air-dried, weighed, and then spiked with isotopically-labelled surrogates. Surrogates should be allowed to equilibrate with the sediments, preferably overnight prior to extraction, to ensure that surrogate recoveries are more representative of the sample matrix.

7.7.3 Chlorinated Phenols, Guaiacols, and Catechols

Procedures for the extraction and derivatization of chlorinated phenolics in effluents are recommended in Table 7.5. Surrogates which should be added prior to extraction are 2-fluorophenol, d6-phenol, 2,4,6-tribromophenol. Internal standards should be d4-1,4-dichlorobenzene, and hexachlorobenzene. Characteristic ions are presented in the work by Lee *et al.* (1989).

7.8 Chemical Tracers

Mills are required where practical to provide confirmation at the time of field sampling that the samples collected are representative of effluent exposed and reference areas. In hydrologically dynamic receiving environments, or those receiving multiple discharges, it will likely be necessary to select a tracer which will be accumulated in fish tissue. The purpose of using a tracer is to verify the exposure of fish to effluent in the near- and far-field areas, and to verify the lack of exposure at reference areas.

This section provides guidance on the selection of tracers in fish, sampling, analysis and QA/QC procedures based on review of Cycle 1 data. A summary of the Tracer Expert

Working Group's conclusions (Tracer Expert Working Group 1997) is also presented at the end of this section.

7.8.1 Applicability of Tracers

Based on the review of the Cycle 1 data, the application of tracers in the adult fish survey is accepted in principle. It is recommended that site-specific decision making processes be followed to provide guidance for identifying mills where the probability of achieving the desired goal is high. Benefiting from the experiences gained in Cycle 1 at the national level, site specificity must be a primary consideration.

The decision tree provided in this section (Figure 7.2) should be followed to determine the applicability of fish tracers in subsequent cycles of EEM. The decision-making process considers factors such as whether fish at the study site can be captured in a local zone with some stability in effluent concentrations, if the mill under study uses a high percentage of softwood furnish, whether barriers separate reference and study sites and the effluent concentrations of resin acids.

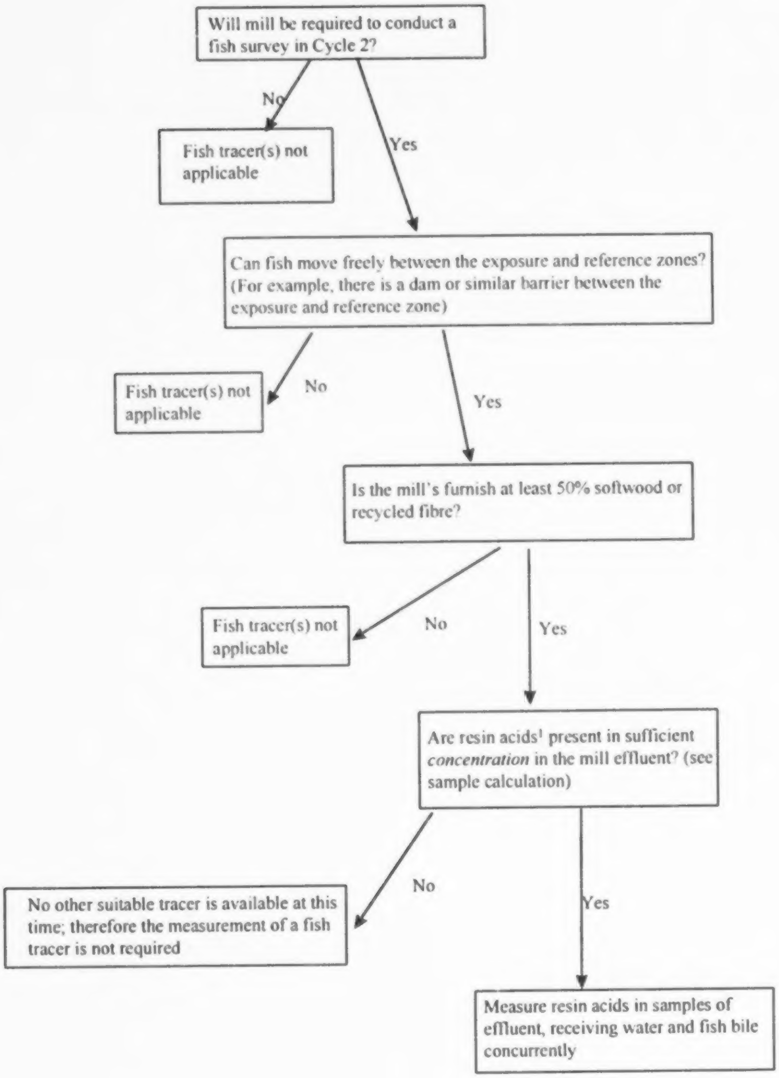
7.8.2 Selection of Tracers

The selected tracers have to be specific for the effluent in question, must be taken up by the fish fairly rapidly, and retained from several days to several weeks in their original or only slightly modified form. Reliable analytical techniques for these compounds must be available and the compounds must be present in concentrations ranging from ng/g (ppb) to mg/g (ppm) to avoid very large errors normally encountered at lower concentrations.

Selection of the candidate chemical or biochemical tracer will be site-specific and depend upon several factors:

- the type of survey being conducted
- the species of interest
- the type of receiving environment (e.g. marine, freshwater, riverine, coastal)
- the characteristics of the effluent (e.g. process and treatment system)
- the presence of physical or biological constraints to movement
- the presence of other discharges to the receiving environment. For waters receiving discharges from many industries, a tracer specific to the effluent under examination must be selected.

Figure 7.2 Fish tracer decision tree


Will mill be required to conduct a fish survey in Cycle 2?
No
Fish tracer(s) not applicable
Yes
Can fish move freely between the exposure and reference zones? (For example, there is a dam or similar barrier between the exposure and reference zone)
No
Fish tracer(s) not applicable
Yes
Is the mill's furnish at least 50% softwood or recycled fibre?
No
Fish tracer(s) not applicable
Yes
Are resin acids¹ present in sufficient *concentration* in the mill effluent? (see sample calculation)
No
No other suitable tracer is available at this time; therefore the measurement of a fish tracer is not required
Yes
Measure resin acids in samples of effluent, receiving water and fish bile concurrently

Note 1 : Other tracers maybe substituted if proven to be effective

Sample Calculation: To determine if the mill effluent contains sufficient concentrations of resin acid to use as a tracer.

Assumptions:

1) Detection limit for resin acids in fish bile = 0.5 ug/g
2) Detection limit for resin acid in water samples = 25 ug/L
3) Bioconcentration factor (BCF) of resin acids in fish bile = 1000
4) Exposure zone fish are captured within the 1% effluent plume

Equations: Tissue concentration = water concentration x BCF
Water concentration = effluent concentration x 0.01

Therefore: Effluent concentration = tissue concentration / (0.01 x BCF)

Effluent Concentration = 0.5 ug/g / (0.01 x 1000)
= 0.05 ug/g
= 50 ug/L

Therefore if the effluent consistently contains 50 ug/L of resin acid, then there should be detectable concentrations of resin acid in the fish bile. This is a very conservative estimate. Detection limits for resin acids in bile can be as low as 0.1 ug/g and BCF for resin acids in fish bile have been reported in the 10

While the choice of a particular tracer may vary by site, some general characteristics apply. A tracer should:

- not be rapidly degraded in the environment or depurated in biota
- be detectable in site effluent at a sufficient concentration to predict detectable levels in the receiving environment, for measurements in the water column
- be detectable in site effluent or, alternatively, in sediments known to have been historically contaminated by site effluent for measurements in fish tissue or bile
- should be unique to the effluent in receiving environments receiving multiple discharges.

In simple riverine systems, where the biota are confined by physical or biological constraints (e.g. dams, sessile behaviour), a chemical tracer sampled from the water column at sampling sites, along with corresponding effluent samples, would likely be sufficient. However, many receiving environments will be considerably more complex than this and a different approach will be needed.

For certain industrial processes, it is possible that no products from the site effluent will accumulate in fish tissue to detectable levels. In these situations, it will only be possible to determine the relative position of the effluent plume at the time of field sampling, using a conservative tracer in the water column.

Elevated concentrations of a tracer with a relative short biological half-life could suggest fish were exposed to the effluent. No defensible statement may be possible on the duration of time that the fish were in that zone because of the probability that a free swimming fish could pass through many concentration gradients. Indices such as fecundity and egg diameters would require many weeks to develop. In cases where fish are collected at a study site where tracer concentrations may be too low to serve its intended purpose, the significance of tracers is greatly reduced.

Another issue to be resolved is clarification on the significance of tracers found in fish from the reference site. Resin acids, fatty acids, phenols, and other natural chemicals are found in wood products; therefore, their ubiquitous distribution in fish from reference sites would not be unexpected. Chlorinated dioxins and furans are industrial by-products and can be distributed by long range atmospheric transport; therefore, their presence in fish from reference sites could be expected.

Resin acids have been identified as a useful tracer in fish in some cases, but other tracers may be substituted if proven to be effective. Resin acids appear to be the most promising tracer for softwood mills; they are by-products of mills processing softwood and aspen, and cannot be tracers for mills whose furnish is primarily other hardwoods.

Dehydroabietic and abietic acids were the two most common tracers identified based on their presence in white sucker. Also, tracers in other media (e.g. sediment) may be useful as part of site-specific monitoring studies. No potential tracers have been identified for hardwood-based mills and further work is needed. It is also necessary to search for potential tracers that survive various forms of secondary treatment of the effluents. Data to date indicate that concentrations of resin and fatty acids and many of the chlorophenols were significantly reduced by secondary treatment or the reduction in the use of elemental chlorine in bleaching (> 80% for resin acids).

Mills considering resin acids as tracers are encouraged to identify laboratories capable of doing such analyses early in the design of the EEM study. Because only a limited number of laboratories are currently capable of doing these analyses, there could be shortages of analytical capacity and/or scheduling difficulties encountered as a result. Additionally, some laboratories may not have experience in these analyses and will require method development time to ensure adequate QA/QC. These factors, and the availability of alternative tracers, should all be considered in making the final selection of a tracer.

The use of dioxins and furans as tracers should be discontinued. It must also be recognized that current mill practices have virtually eliminated, the discharge of dioxins and furans in effluents. Residual amounts that may be detected in fish are likely attributable to sources deposited in sediments from previous mill practices rather than the quality of the mill effluent presently discharged.

Fatty acids cannot be used as tracers, mainly because of their common presence in biological substrates.

A paucity of data from mills which already conducted Cycle 1 prevented detailed evaluation of the suitability of using chlorophenols as tracers.

In cases where there are no suitable compounds in the effluent that could serve as a tracer in fish, a mill could use a water column tracer approach. There are a broader range of compounds or measurements (e.g. conductivity, colour, sodium) that can be measured in the receiving waters to confirm and quantify the presence of the mill effluent. This approach could at least confirm that the fish were captured in an area where there was mill effluent present. The weakness of this approach lies in not knowing anything about the duration of fish exposure to the effluent.

7.8.3 Sample Collection and Analysis

Samples must be collected in accordance with the quality assurance/quality control principles identified in Chapter 4 and in accordance with the relevant performance criteria identified in Annex 1, where applicable.

Many fish species are very mobile and can move between exposed and reference sites during the course of the year. Consequently, indicators of individual organisms may be needed to predict individual exposure. Species selection is determined by availability, and factors such as low mobility and prolonged residency time to the effluent exposure are important.

Fish or shellfish species must be readily available in the study and reference zones. Sampling protocol, such as the number of specimens taken, their size (age) and sex, and time of the year must be established. Information must be provided for sample storage and processing. It should be decided whether to analyze individual specimens or pooled samples and whether the analyses are to be done singly or in duplicate, and how to deal with unusual values.

The portion of fish sampled for a chemical tracer depends upon the industrial process of the site and the degree of effluent treatment. For example, for pulp mills (chlorine and non-chlorine bleaching) equipped with primary treatment, the preferred option is to measure the level of chemical tracer in tissue (e.g. muscle, liver, hepatopancreas) since this is representative of long-term exposure. For mills having secondary treatment, it may be necessary to determine tracer levels in bile, as levels in edible tissues may be below detection limits.

The use of single composite samples must be eliminated for tracer purposes to provide a statistical basis for data evaluation and decision making. Future sampling should include multiple individual samples or multiple pooled (composite) samples.

Fish should be analyzed for tracers if a practical method is available. It is recommended that 10-20% of fish collected be sampled. It should be noted that pulp mills which use, or have used, chlorine bleaching must measure tissue levels of chlorinated dioxin and furan congeners on edible portions of fish (Section 9.1, Annex I).

It is recommended that bile be analyzed for resin acids and their conjugates. Resin acids are magnified 1,000 to 100,000 times in bile and so make very good tracers. For this analysis, it is suggested that the technique described in Morales *et al.* (1992) be used. The following acids may be monitored: pimaric, sandaracopimaric, isopimaric, dehydroabietic, and abietic. Analysis for fatty acids should be avoided since they are common contaminants and are associated with soaps which are used by laboratories in glassware washing. There is the possibility that analyzing for fatty acids may lead to false positives.

A QA/QC program for the measurement of tracers in samples of fish tissue is required. The analytical methods, and quality assurance/quality control measures, used for chlorinated dioxins and furans, resin and fatty acids, and chlorinated phenols are described in Table 7.6. The participation in round robin testing particularly for resin acids, chlorinated phenols, and other tracers that may be considered in the future, must be a routine practice, similar to those available for laboratories that analyze for dioxins and furans. Certified reference materials (CRMs) may not be available for many of these chemicals and, therefore, must be developed and become an integral part of the analytical validation process.

In terms of quality control for bile analysis, both resin acid and chlorophenol analyses of bile require hydrolysis of the bile prior to extraction, derivatization, and analysis. Either enzymatic or acid hydrolysis is acceptable (Söderström *et al.* 1994). A laboratory must show that the hydrolysis is effective, otherwise data will be reported which may be misleading. One suggested method is fortifying bile samples with either _-naphthyl-_-D-glucuronic acid or 6-bromo-2-naphyhyl-_-D-glucuronic acid. After hydrolysis, the recovery of _-naphthol and/or 6-bromo-2-naphthol is reported for each sample. This provides quality control data on the efficiency of the hydrolysis as well as for the efficiency of extraction and derivatization. Without this type of data the results are largely useless.

For quality control measures in resin acids and chlorophenols in bile, the lab should refer to Morales *et al.* (1992).

7.8.4 Tracers in Benthos

Cycle One data showed that sediment chlorinated guaiacols and catechols provided the best tracer information for the bleached kraft mills. After secondary treatment has been installed at a mill, the use of a particular tracer may need to be re-evaluated and secondary treatment may affect the usefulness of certain tracers, particularly resin acids.

Where benthic studies are proposed, details should be discussed with the regional EEM contacts.

7.9 References

Alberta Environment. 1990. Method AE 129.0 Regional Fatty Acids in Pulp Mill - Effluent and Receiving Waters.

Alberta Environment. 1991. Method AE 130.0 Chlorinated Phenotic Compounds in Bleached Kraft Mill Effluents and Receiving Water.

American Soils Association. 1982. Methods of Soil Analysis Properties. 2nd ed. American Society of Agronomy Inc., Soil Science Society of America Inc. American Society of Agronomy and Academic Press, Madison Wisconsin.

APHA. 1989. Standard Methods for the Examination of Water and Wastewater. 17th ed. American Public Health Association Washington, D.C.

APHA. 1992. Standard Methods for Examination of Water and Wastewater. 18th edition, published by the American Public Health Association (APHA).

APHA. 1995. Standard Methods for Examination of Water and Wastewater. 19th edition, published by the American Public Health Association (APHA).

ASTM. 1986. Manual on Presentation of Data and Control Chart Analysis. Committee E-11 on Statistical Methods. ASTM Special Technical Publication 15D. American Society for Testing and Materials, Philadelphia.

Birkholz D.A., R.T. Coutts and S.E. Hrudey. 1988. Determination of polycyclic aromatic hydrocarbons in fish tissues. J. Chromatogr. 449:251.

British Columbia Ministry of Environment. 1991. Water Quality Branch Objectives in 1990. Victoria, BC.

Burd, B.J. 1997. British Columbia Salmon Aquaculture Review: Discussion Paper on Waste Discharges. Prepared for the Environmental Assessment Office SAO, May 1997. 181pp.

CAEAL (Canadian Association of Environmental Analytical Laboratories). 1991. Code of Practice and QA Manual for Laboratory Analysis of Sewage Treatment Effluent in Support of the MISA Program; Draft report prepared for CAEAL and the Ontario Ministry of the Environment by Zenon Environmental Laboratories.

Calvert, S. E., T.F. Pedersen, P.D. Naidu and U. von Stackelberg. 1995. On the organic carbon maximum on the continental slope of the eastern Arabian Sea. Journal of Marine Research 53: 269-296.

Canadian Council of Ministers for the Environment (CCME) 1995. Protocol for the derivations of canadian sediment quality guidelines for the protection of aquatic life. Report CCME EPC-98E. 38pp.

Carron, J.M. and Afghan, B.K. 1989. Environmental Aspects and Analysis of Phenols in the Aquatic Environment. Analysis of Trace Organics in the Aquatic Environment, CRC Press.

ELAP (Environmental Laboratory Approval Program). 1988. Environmental Laboratory Approval Program Certification Manual. New York State Department of Health.

Environment Canada. 1992a. Reference Method for the Determination of Polychlorinated Dibenzo-para-dioxins (PCDDs) and Polychlorinated Dibenzofurans (PCDFs) in Pulp and Paper Mill Effluents. Report EPS 1/RM/19. 45 pp.

Environment Canada. 1992b. Internal Quality Assurance Requirements for the Analysis of Dioxins in Environmental Samples. Report EPS 1/RM/23. 26 pp.

Hargarve, B.T., Duplisea, D.E., Pfieffer, E. and Wildish, D.J 1993. Seasonal changes in benthic fluexes of dissolved oxygen and ammonia associated with marine cultured Atlantic salmon. Mar. Ecol. Prog. Ser. 96: 249-257.

Hargrave B.T.,G.A.Phillips,L.I.Doucette,M.J.White,T.G. Milligan, D.J.Wildish and R.E.Cranston. 1997. Assessing benthic impacts of organic enrichment from marine aquaculture.Wat.Air Soil Pollut.(In press).

Hargrave,B.T.,G.A.Phillips,L.I.Doucette,M.J.White,T.G.Milligan,D.J.Wildish and R.E. Cranston. 1995.Biogeochemical observations to assess benthic impacts of organic enrichment from marine aquaculture in the Western Isles region of the Bay of Fundy, 1994. Can. Tech. Rep. Fish. Aquat. Sci. No. 2062. 159pp.

Holme, N.A. and A.D. McIntyre (Eds). 1984. Methods for the Study of Marine Benthos, Second Ed., Blackwell, Oxford, 387pp.

Hunt, D.T.E. and A.L. Wilson. 1986. The Chemical Analysis of Water; General Principles and Techniques, 2nd Ed. The Royal Society of Chemistry, London.

Hynning P.-A., M. Remberger and A.H. Nielson. 1989. Synthesis, gas-liquid chromatographic analysis and gas chromatographic - mass spectrometric identification of nitrovanilline, chloronitrovanillins and chloronitroguaiacols. J. Chromatogr. 467:99.

King, D.E. 1976. Quality Control and Data Evaluation Procedures. Section I. Analytical Responsibility. Special Report to Laboratory Services Branch, Ontario Ministry of the Environment.

King, D.E. 1989. Code of Practice for Environmental Laboratories Special Report to the Ontario Ministry of the Environment, ISBN 0-7729-5874-2.

Klemm, D.J., P.A. Lewis, F. Fulk and J.M. Lazorchak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. EPA/600/4-90/030.

Lee, H.B., R.L. Hong-Yow and P.J.A. Fowlie. 1989. Chemical Derivatization Analysis of Phenols. Part VI. Determination of Chlorinated Phenolics in Pulp and Paper Effluents. J.A.O.A.C. 72(6), 979-984.

Macdonald, R.W., R.M Macdonald, M.C. O'Brien and C. Gobeil. 1991. Accumulation of heavy metals (PB, Zn, Cu, Cd), carbon and nitrogen in sediment from Strait of Georgia, B.C. Canada. Marine Chemistry 34:109-135.

Morales, A., D.A. Birkholz, and S.E. Hrudey. 1992. Analysis of pulp mill effluent contaminants in water, sediment, and fish bile - fatty and resin acids. Water Environment Research. 64:660-668.

Nagpal, N.K. and Pommen, L.W. 1993. Approved and Working Criteria for Water Quality - 1993. Rep US Ministry of Lands, Parks and Envir., November 1993. 44pp.

NCASI (National Council of the Paper Industry for Air and Stream Improvement Inc.). 1986. Procedures for the Analysis of Resin and Fatty Acids in Pulp Mill Effluents. Technical Bulletin No. 501, Washington.

NLET (National Laboratory for Environmental Testing) 1994a. Manual of Analytical Methods, Vol.1. Major Ions and Nutrients. NLET, NWRI, Environment Canada, Burlington, Ontario.

NLET (National Laboratory for Environmental Testing) 1994b. Manual of Analytical Methods, Vol.2. Trace Metals. NLET, NWRI, Environment Canada, Burlington, Ontario.

NLET (National Laboratory for Environmental Testing) 1996. Manual of Analytical

Methods, Vol.3. Trace Organics. NLET, NWRI, Environment Canada, Burlington, Ontario.

OMOE. 1988. Estimation of Analytical Detection Limits (MDL) Report by the Ontario Ministry of the Environment ISBN-0-7729-4117-3.

OMOE. 1989. The Determination of Resin And Fatty Acids in Effluents and Water by GLC-FID. PWRFA-E3166A, Laboratory Services Branch, Rexdale, Ontario.

Poole, N.J., D.J.Wildish and D.D. Kristmanson. 1978. The effects of the pulp and paper industry on the aquatic environment. CRC Crit.Rev. Environ.Control 8:153-195.

Söderström, M., C.A. Wachmeister and L. Förlin. 1994. Analysis of chlorophenolics from bleach kraft mill effluents (BKME) in bile of perch (Perca fluviatilis) from the Baltic Sea and development of an analytical procedure also measuring chlorocatechols. Chemosphere. 28:1701-1719.

Striplin Environmental Consultants Inc. 1996. Development of reference value ranges for benthic infauna assessment endpoints in Puget Sound. Report to Wash.Dept. Ecology, Sediment Management Unit, Jan30, 1996.

Sundby, B. 1974. Distribution and transport of suspended particulate matter in the Gulf of St. Lawrence. Can. J. Earth Sci. 11: 1517-1533.

Taylor, J.K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers Inc., Chelsea, Michigan.

Tetra Tech Inc. 1986. Recommended protocols for measuring conventional sediment variables in Puget Sound. Prepared for the U.S. EPA Region 10. Final Report TC-3991-04 (March 1986).

Tetra Tech, Inc. 1987. Tetra Tech, Inc. 1987. Recommended Protocols for Sampling and Analyzing Subtidal Benthic Macroinvertebrate Assemblages in Puget Sound. A Report prepared for U.S. EPA, Region 10, Seattle. Final Report TC-3991-04.

Tracer Expert Working Group. 1997. Tracer Expert Working Group Report. Prepared for the EEM Technical Management Committee. EEM/1997/2.

U.S. EPA. 1984. Definition and Procedure for the Determination of the Method Detection Limit - Revision 1.11; Appendix B to Part 136. Federal Register Vol. 49, No. 209, Oct. 26,1984, Part VI, 40 CFR Part 136.

U.S. EPA. 1986. Test Methods for Evaluating Solid Waste. Physical/Chemical Methods. Publication No. SW-846, 3rd Edition. Washington, D.C. pp 9060 1-4.

U.S. EPA. 1988. User's Guide to the Contract Laboratory Program. CLP Sample Management Office.

U.S. EPA. 1990a. Method 1613: Tetra through Octa-chlorinated Dioxins and Furans by Isotope Dilution HREC/HRMS.

U.S. EPA. 1990b. Analytical Procedures and Quality Assurance Plan for Determination of PCBP/PCDP in fish EPA/600/3-90/022.

U.S. EPA. 1990c. Proposed Glossary of Quality Assurance Related Terms. QAMS RD-680. Draft report.

U.S. EPA 1991. Method 1653: Chlorinated Phenols in Wastewater by in situ Acetylation and GC/MC.

Voss, R.H. and A. Rapsomatiotis. 1985. An improved solvent extraction based procedure for the gas chromatographic analysis of resin and fatty acids in pulp mill effluents. J. Chromatogr. 346: 205-217.

Walton, A. 1978. Methods for Sampling and Analysis of Marine Sediments and Dredged Materials. Ocean Dumping Report 1, Department of Fisheries and Environment, Ottawa.

Wildish,D.J.,J.L.Martin, R.W.Trites and A.M.Saulnier 1990.A proposal for environmental research and monitoring of organic pollution caused by salmonid mariculture in the Bay of Fundy.Can.Tech.Rep. Fish.Aquat. Sc. No.1724,24 pp.

8 Process Effluent Sublethal Toxicological Testing

8.1 Background

Toxicity testing with one or more acute lethality tests has been required for some years by federal or provincial regulations for pulp and paper mills. Mills that comply with non-lethality requirements of regulations would not be expected to produce acute toxic effects in the receiving water body. Consequently, the Environmental Effects Monitoring (EEM) provisions of the 1992 Pulp and Paper Effluent Regulations focus on assessing whether or not fish populations, fish habitat and use of fisheries resources are being protected at specific effluent discharge sites. One component of EEM is the measurement of potential sublethal effects through the use of toxicity tests. The three specific objectives for the use of sublethal toxicity testing in EEM are: 1) to contribute to the field program as part of the weight of evidence approach; 2) to compare process effluent quality between mill types and measure changes in effluent quality as a result of effluent treatment and process changes; and 3) to contribute to the understanding of the relative contributions of the mill in multiple discharge situations. Further details are provided in the report of the EWG on Laboratory Toxicology (Environment Canada 1997a).

The required sublethal toxicity tests are described in Table 8.1.

8.2 Tabulation of Endpoints and Validation of Test Results

The geometric mean of all IC25s (GM-IC25) (EC25s and LC50s are also used but IC25 is discussed here for illustrative purposes) for a given species should be calculated for each cycle. These results can be used to track changes in effluent quality between cycles and may also be useful in contributing towards a weight of evidence approach. The GM-IC25 results can also be used to increase understanding of the relative contribution of the mill effluent in multiple discharge situations.

Table 8.1 Methods for Toxicity Tests with Pulp and Paper Process Effluents (Environment Canada 1997b)

Test Description	Receiving Environment	Test Species	Acceptable Method(s)
Fish early life stage development test	Marine	*Menidia beryllina* (inland silverside) or (*Atherinops affinis)* topsmelt	U.S. EPA 1994 and 1995
	Freshwater	*Pimephales promelas* (fathead minnow) or salmonid spp.*	Environment Canada 1992a or/ Environment Canada 1997c
Invertebrate reproduction test	Marine	Echinoids (sea urchins or sand dollars)	Environment Canada 1992b
	Freshwater	*Ceriodaphnia dubia*	Environment Canada 1992c
Plant toxicity test	Marine	*Champia parvula*	U.S. EPA 1994
	Freshwater	*Selenastrum capricornutum*	Environment Canada 1992d

*For Canadian receiving environment locations, west of the Rocky Mountains, where fathead minnows are not an indigenous species, a salmonid species must be used according to Environment Canada (1997c).

8.2.1 Validation of Test Results

Procedures for quality assurance and quality control (QA/QC) must be followed by both the field crews collecting environmental samples and the laboratory carrying out the toxicity testing, as discussed in Chapter 4 of this document and the required toxicity test method documents. These measures optimize accuracy and precision within the limitations of the test method and natural biological variability.

Therefore, a first step in the interpretation of toxicity data for EEM should be the resolution of any problems with QA/QC. The following specific QA/QC practices and recommendations have been incorporated into Annex 1 as requirements (Environment Canada 1997a, 1997b).

- sublethal toxicity tests that fail to meet test method validity criteria must be repeated on a new sample;
- compliance to effluent sample age limit or re-sampling required (i.e. test within 3 days of sample);
- adherence to minimum level of reporting required by the test method (see

example on Table 8.2);

- reporting of test data within 90 days of test completion for quality assurance check;
- reporting of "less than" values as a test endpoint will no longer be acceptable;
- all test endpoints must be bracketed by at least one test concentration;
- a reference toxicant test must be conducted within 30 days of the effluent test; and
- reference toxicant tests must be performed under the same experimental conditions as the effluent test.

These QA/QC changes focus on ensuring that valid toxicity data of high quality will be generated by private testing laboratories. Similarly, data could be declared invalid if one or more essential elements of the test method were not followed (e.g. failure to meet organism health criteria, inappropriate manipulations of the sample, failure to conduct the minimum in-test monitoring, incorrect statistic used for endpoint calculation, etc).

Toxicity data submitted as part of the EEM program for the pulp and paper industry must be accompanied by description of the materials, methods, and calculations for each test. Minimum reporting requirements are detailed in Section 8 or 9 of the toxicity test method documents of Environment Canada. An example of minimum reporting requirements for the *Ceriodaphnia* test is outlined in Table 8.2. In addition, test report checklists have been developed for assessing the validity of test results for each of the 8 test options which will be distributed to all mills and EEM Regional Authorization Officers. The minimum reporting requirements for methods of U.S. EPA (1994a and 1995) have been prescribed for the purpose of subsequent EEM cycles. The U.S. EPA requirements generally conform to Environment Canada specifications.

Laboratories contracted by the pulp and paper industry to conduct sublethal toxicity testing should be accredited by the Standards Council of Canada through the Canadian Association for Environmental Analytical Laboratories (CAEAL), the Ministère de l'Environnement et de la Faune (MEF) du Québec or an organization offering an equivalent level of accreditation. These accreditation organizations follow International Standards Organization (ISO) standards which focus on the quality system and practices of a laboratory. This level of quality assurance should be a prerequisite for selecting a testing laboratory.

Table 8.2 Minimum Reporting Requirements for *Ceriodaphnia dubia* Sublethal Toxicity Test

Effluent
Sample type (eg: whole effluent, final effluent etc.)
Information on labelling/coding of sample
Temperature of sample upon arrival at laboratory
Date for sample collection and date/time for sample receipt at laboratory
Date for test start

Test Organisms
Species
Source
Age at start of test
Any unusual appearance, behaviour, or treatment of test organisms, before their use in the test

Health Criteria:
➢Mean % mortality of the brood organisms in individual cultures during 7-d period preceding test
➢Mean # of surviving young produced per adult in individual cultures during 7-d period preceding test
➢# of young produced by each brood organism in its previous brood
➢Ephippia in the culture

Test Conditions and Facilities
Test method
Dates or days during test when sub-samples or multiple samples used
Date for test completion
Description of test vessels (size, shape, type of material)
Person performing test
Rate of pre-aeration of solutions, if required, prior to addition of test organisms to the test vessel
Duration of pre-aeration
Rate/duration of aeration, if any, during exposure of test organisms
Procedure, if any, for pH adjustment of sample
Procedure, if any, for sample filtration or hardness adjustment
Type(s) and source(s) of control/dilution water
Type and quantity of any chemical(s) added to control/dilution water
# and concentrations of test solutions
Volume and depth of solution in each test vessel
# of replicates per concentration
# of organisms
D.O. and pH of sample just before its use
Temperature, D.O. and pH of test solutions and controls
Anything unusual about the test, any deviation from the test method, any problems encountered, any remedial measures taken

Results
Mean % mortality for first-generation daphnids in each test solution (including control) at each period of observation and test end
Mean # of neonates per first-generation daphnid in each test solution (including control), as noted during each observation period and test end
LC50 (and 95% confidence limits) for survival; indication of quantal statistic employed
IC25 (and 95% confidence limits) for reproduction; indication of quantitative statistic used
Results and duration of the reference toxicant test(s)
Was a reference toxicant test conducted within 30 d of the *Ceriodaphnia* test being performed on an effluent?
Was it conducted under the same experimental conditions?
Were the test endpoints bracketed by at least 1 test conc. (except for "> 100%")?
Was report submitted within 90 d of test completion ?

8.2.2 Tabulation of Endpoints

Valid sublethal toxicity test results should be available for the three test species (plant, invertebrate, and fish) from each of the sublethal toxicity test reports to be submitted. The IC25 (or EC25 for the salmonid embryo test) estimate of toxicity should be extracted from each test report. If the effluent does not cause a 25% sublethal inhibition or effect for any of the freshwater sublethal tests, then an IC25/EC25 cannot be calculated and it is reported as >100%. For the estuarine/marine sublethal tests, the highest possible test concentration will be 67% effluent due to the addition of hypersaline brine. Therefore, an IC25 of >67% would be reported for estuarine/marine tests which show little or no effect at the maximum concentration. For the purpose of Step 3 calculations in this situation, the IC25 or EC25 should be set equal to 100% or 67% for freshwater and estuarine/marine tests, respectively[1].

Mortality in some of the concentrations might also prevent calculation of the IC25/EC25. For example, there might be no measured effects (mortality, growth or reproduction) in 32% concentration, but appreciable mortality in 56%. It would then be impossible to obtain a good estimate of the inhibition of growth or reproduction for the 56% concentration, and hence impossible to determine an IC25/EC25. In such a situation, the IC25/EC25 should be assumed to be equal to the higher concentration, which in this case is 56%.

There should be IC25s for each of the three test species (EC25 for salmonid embryo test). This information is summarized by calculating the geometric mean for each set of IC25s[2]. For example, testing of a mill effluent on six occasions resulting in measured IC25 values of 10, 15, 17, 23, 25, and 30% would lead to a geometric mean of 19%.

8.3 Data Interpretation in Relation to the Toxicity Objectives

8.3.1 Changes in Effluent Quality

The geometric mean of the IC25 (GM-IC25) for each species can be compared between cycles to assess changes in the quality of effluent over time at each mill or between mills and mill types. Improvements are expected as a result of changes in process or effluent treatment

---

[1] Although this action might seem to be biased against the industry being evaluated, that is really not the case. Assigning a value of 100% (or 67%) will almost always mean that this hypothetical IC25 (or EC25) is closer to the other measured IC25s (or EC25), reducing the standard deviation.

[2] The geometric mean may be calculated as the nth root of n numbers multiplied together. Alternatively, the logarithms of the n IC25s (EC25s or LC50s) may be added together, the sum divided by n, and the antilog of the result is the geometric mean.

8.3.2 Understanding Multiple Discharge Situations

Comparison of mill effluent toxicity data to sublethal toxicity data from other nearby industrial and/or municipal discharges, can help in understanding the relative contribution of the potential impact of each effluent source on the environment. Provided the relevant toxicity, flow and dispersion information is available for the other discharges, the inter-relation or overlap of the effluent fields may be better understood. Another method of comparing relative loading (toxic contribution) from each source is to calculate the TER [toxicity emission rate = (100/GM-IC25) x flow]. However, this calculation does not relate to the receiving environment since the effluent dispersion and dilution have not been considered.

8.3.3 Contribution to the Weight of Evidence Approach

The recommended approach has been modified from procedures of the U.S. EPA (1991). In order to relate the sublethal toxicity test results to the receiving environment, the GM-IC25 for each species needs to be considered in relation to the dilution available in the discharge environment. As Annex 1 requires that each EEM study delineate the extent of the plume to the 1% effluent isopleth, the potential zone of effect can be estimated for each mill by dividing the distance of the 1% zone by the geometric mean of the IC25 (Table 8.3). This potential effects zone can be estimated as that area of the plume in which concentrations exceed the geometric mean of the IC25s. The comparison could be a simple narrative statement relating plume concentrations to the geometric mean of the IC25s. Therefore, the potential effects zone describes the maximum distance where a specified effect might be found for a specific species.

A potential effects zone may be interpreted as a rough indication of the extent of 25% inhibition by effluent in the environment. If, for example, a mill's GM-IC25 for a particular test was 1% v/v, this would match the extent of the 1% zone. Assuming that a mill's effluent concentration decreases with distance from the outfall, if a mill's GM-IC25 was 50% v/v, the extent of 25% inhibition in the environment would be 1/50 the 1% zone. Fish, invertebrate, and plant IC25s are not expected to be similar due to differing species sensitivities and test method endpoints.

Table 8.3 demonstrates how the zone of potential sublethal effect could be reduced as sublethal toxicity was reduced by the installation of secondary treatment between Cycle 1 and 2.

Table 8.3 Hypothetical Example of Potential Zone of Effect Estimates (Assuming a 1% Zone of 1500 m).

Species	GM IC25(Cycle1)	GM IC25(Cycle2)	Potential Zone of Sublethal Effect	
			Cycle 1	*Cycle 2*
Selenastrum sp.	18 %	36 %	83 m	42 m
Ceriodaphnia sp.	6 %	12 %	250 m	125 m
Fathead minnow	90 %	>100 %	17 m	<15 m

8.3.4 Considerations for Integration of Toxicity Test Results

Certain points should be considered when toxicity data are used to estimate a potential effects zone.

(1) *Laboratory Results*: Sublethal laboratory tests provide estimates of toxicity under strictly controlled laboratory conditions for each test species. These conditions do not replicate environmental conditions at the site under study.

(2) Species differences: Species differences in sensitivity to pulp and paper effluents must be taken into account when extrapolating results from laboratory sublethal toxicity tests to effects on indigenous biota.

(3) *Background toxicity*: The description above assumed that there was no other upstream contributions of toxicity. That assumption would be erroneous if there were overlapping plumes.

(4) *Type of receiving water*: Receiving water pH, hardness, DOC and other modifying factors could potentially increase or decrease toxicity of the effluent compared to tests with laboratory water.

(5) *Plume uncertainties*: Calculations of dilution might be difficult or inaccurate, or the position of the mixing zone might be variable. Where this uncertainty exists, estimating a zone of potential effect would have an equal level of uncertainty.

8.4 References

Environment Canada. 1992a. Test of reproduction and survival using the cladoceran *Ceriodaphnia dubia*. Environment Canada, Ottawa, Ontario. Report EPS 1/RM/21.

Environment Canada. 1992b. Test of larval growth and survival using fathead minnow. Environment Canada, Ottawa, Ontario. Report EPS 1/RM/22.

Environment Canada. 1992c. Growth inhibition test using the freshwater alga *Selenastrum capricornutum*. Environment Canada, Ottawa, Ontario. Report EPS 1/RM/25.

Environment Canada. 1992d. Fertilization assay with echinoids (sea urchins and sand dollars). Environment Canada, Ottawa, Ontario. Report EPS 1/RM/27.

Environment Canada. 1997a. Report of the Expert Working Group on Laboratory Toxicology, EEM Technical Management Committee. EEM/1997/7.

Environment Canada. 1997b. Pulp and Paper Aquatic Environmental Effects Monitoring Requirements, Environment Canada. Annex 1 to EEM/1997/1.

Environment Canada. 1997c. Toxicity tests using early life stages of salmonid fish (rainbow trout, coho salmon or Atlantic salmon), Second Edition. Environment Canada, Ottawa, Ontario. Report EPS 1/RM/28.

U.S. EPA (United States Environmental Protection Agency). 1991. Technical support document for water quality-based toxics control. Environmental Protection Agency, Office of Water, Washington, D.C. EPA/505/2-90-001.

U.S. EPA (United States Environmental Protection Agency). 1994a. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and freshwater organisms. (Second edition). Environmental Monitoring Systems Laboratory, Cincinnati, Ohio. EPA/ 600/4-91-003.

U.S. EPA (United States Environmental Protection Agency). 1995. Short-term methods for estimating the chronic toxicity of effluent and receiving waters to west coast marine and estuarine organisms, First edition. U.S. Environmental Protection Agency, Cincinnati, Ohio.

---

9 Recommended Outline for Interpretative Reports

9.1 Introduction

The following outline was developed for mills to use when preparing their interpretative reports. It is based on Annex 1 for Aquatic Environmental Effects Monitoring of Pulp and Paper Mills (Environment Canada, 1997) and this Technical Guidance Document with respect to design and reporting requirements. It is also the basis for information required for the electronic reporting spreadsheets (ERS). This outline will promote national consistency in reporting results.

9.2 Information Requirements

9.2.1 Site Description and Mill Update

To ensure that each EEM study is evaluated objectively, it is essential that a synopsis of each study with regards to mill history and operations, ecological aspects and study design be included in the report. It is particularly important to include any changes or variations to the study design, for example, information such as changes which occurred after the study was authorized and before the field work was conducted. Information requirements for this section fall under the following three categories. Other information may be required on a site specific basis.

Synopsis of Mill History and Operations

include:
- significant changes that have been made to the actual mill site (e.g., altered mill process, new effluent treatment and operations)
- was mill in compliance during EEM study? Yes or No (supply relevant history of compliance information with regards to PPER during EEM study period)

Synopsis of Study Area - Ecological Aspects

include:
- any ecological variations that have occurred in the study area since submission of the pre-design report (e.g., new confounding factors)
- any unusual significant events that may have occurred (e.g., floods, spills etc.)
- any other influences (anthropogenic) that may have affected the study area
- any new information that was not available in the pre-design report

Synopsis of Study Design

include:

- any changes from initial study design and rationale for such changes

9.2.2 EEM Study

It is the intention of this section to ensure that all pertinent information related to field and laboratory methodology is given.

The information in this section is generic to all monitoring requirements (Invertebrate Community Survey, Fish Survey, Toxicological Testing, Tissue Analyses, and Tainting Evaluation) and should be provided for each. For specific details on reporting for each monitoring requirement, refer to the respective sections of Annex 1 (Environment Canada, 1997) and this Technical Guidance Document.

The information can be reported under the following categories (other specific information required and important in addition to the synopsis are listed below):

Location

Include any information pertaining to the mill site, study area and reference area.

- comprehensive station location maps (included from study design)
- all station positions (latitude and longitude, datums, station numbers, etc.)
- rationale for choosing both reference and study area
- photographs of both reference and study areas

Methods

Include information on procedures and techniques used to perform the study.

- problems that may have occurred and their solutions
- justification of any variations in methods from that stated in the original study design to meet EEM requirements

Data

Include all raw data, tables, and appendices.

- data organized by taxonomic levels (e.g., order, family, genus, species)
- data in spreadsheet format (hard copy)
- ensure raw data reported in both appendices and in electronic format
- all QA/QC data and assessment (e.g., SRM's, field blanks, calibration charts)

Statistics

Include methodologies, references, assumptions and justifications.

- methods used for statistical analysis including transformations, assumptions, and outliers
- document any data screening
- an independent review of assumptions, approaches, and interpretations
- an assessment of the variability, power analysis, resolution, level of confidence, and the ability to detect a specified change

Results

Include graphics, data summary tables or figures, and summary of statistical analysis.

- all graphics, include interpretation of figures
- any factors affecting the evaluation
- a summary of statistics
- comparison of habitat between reference and study areas
- the effect of outliers or extreme values on the result, if any

Discussion

Include discussion of the observed similarities or differences between the reference and study area, comparison of present results to previous studies in the same area or other pulp mill studies in literature, implications of study results on subsequent cycles, and identification of problems encountered in the monitoring cycle.

- for the Fish Survey and Invertebrate Community Survey only, indicate the type of chemical or biochemical tracer, method used, and justification behind the selection
- whether a difference was evident between reference and study areas

- an evaluation of the results
- the implications of results on subsequent cycles, generally and on a site-specific basis any problems encountered, site-specific solutions/approaches possible for future Cycles
- any implications of the factors affecting the evaluation, power analysis, resolution, trend detection, and implications of the present results on future sampling needs
- comment on the validity of reference areas selected
- discuss the validity and relevance of the methodologies used
- summary of adherence to data quality objectives, standard operating procedures, and identification of any QA/QC problems

Conclusions

In this section of the report, please provide an overall assessment of all the combined results and discussions (this information should not be cause-and-effect type conclusions or "no effect" type conclusions). Include suggestions for modification for EEM to improve program usefulness. The information that is provided here will help in the development of subsequent cycles.

9.3 References

Environment Canada. 1997. Pulp and Paper Aquatic Environmental Effects Monitoring Requirements. Annex 1 to EEM/1997/1

10 Information Management

Data that is collected under the EEM program must be entered into the program's Electronic Reporting Software (ERS) which has been made available to pulp and paper mills. The purpose of the ERS is to ensure a consistent electronic format for uploading into the National EEM Database. Data collected in EEM include physico-chemical, toxicological, biological, mill background, and other site-specific parameters, gathered in cycles spanning several sampling seasons.

The accepted ERS format is Excel-based spreadsheets which are available through Environment Canada's EEM Website (http://www.ec.gc.ca/eem/english/default.htm) or on floppy disk by contacting the National EEM Office (phone: 819 997-1535, fax: 819-953-0461, E-Mail : eem@ec.gc.ca).

The National EEM Database has been created for the archiving and retrieval of information collected during EEM studies across Canada. Primarily, the database will assist Environment Canada and other stakeholders in assessing the adequacy of the pulp and paper regulations:

- checking past trends to determine whether effects on the environment are stable, increasing or decreasing, allowing more accurate prediction of future effects
- comparing effects with similar sites nationwide
- identifying common features and significant differences
- linking with other data sources

Access to the database via the internet will also make EEM data easily available to all interested stakeholders, including all levels of government, industry, the scientific community and the general public.

Although data will be technically validated to a certain extent by government, before acceptance for uploading to the National Database, the pulp and paper mill manager at the source of the data generation is wholly responsible for the quality of the submitted data. This is particularly important for the data generator since there is no automatic data validation scheme built into the spreadsheets. The following steps should be followed in meeting the electronic reporting requirements for pulp and paper EEM :

1. Pulp and paper mills (and consultants) should consult the EEM Website or National EEM Office for the current version of the ERS (data requirements may change in the future as monitoring approaches change) with accompanying detailed instructions for filling in the spreadsheet.
2. Completed spreadsheets must be submitted to Regional Authorization Officers (RAO) according to dates specified in Annex 1. It should be noted that Toxicology results are now reported twice a year (submitted to the RAO within 90 days of testing); all other EEM monitoring elements are reported at the end of the Cycle.
3. Once the RAO has accepted the data, the spreadsheets will be transmitted to the National EEM Office for uploading to the National Database.

GLOSSARY

Batch - A specific number of samples and QC samples of the same matrix processed as a group for analysis of the same variable(s), using the same method.

Benthic (macro) Invertebrates - Those animals without backbones living on and/or in the bottom of a water body (ocean sea, lake, river, stream).

Calibration control - quality control activity designed to demonstrate that the calibration of an analytical run is under statistical control and has not changed significantly between-runs and within a run.

Data Quality Objectives (DQOs) - Pre-defined criteria for the quality of data generated or used in a particular study, so as to ensure that the data are of suitable quality to meet the needs of the program.

Endpoint - The variable (e.g., time, reaction of the organisms, etc,) that indicates the termination of a test, and also means a measurement or value derived that characterizes the results of the test (NOEC, LC_{50}, etc.).

Field sub-sample - A field sub-sample will consist of an individual area or time-limited collection (e.g., one grab, core, cylinder, kick or U-net sample.

Good Laboratory Practices - General guidelines or formal regulations for performing basic laboratory operations or activities that are known to enhance the quality and integrity of the results.

ICp - The percentage inhibition concentration. It represents a point estimate of the concentration of test material that causes a specified percentage reduction in a qualitative biological measurement such as growth of fish. For example, an IC_{25} would be the concentration estimated to cause a 25% reduction in growth of larval fish, relative to the control.

Indigenous - those species native to a particular habitat.

LC50 - The median lethal concentration, i.e., the concentration of material in water that is estimated to be lethal to 50% of the test organisms. The LC50 and its 95% confidence limits are usually derived by statistical analysis of mortalities in several test concentrations after a fixed period of exposure. The duration of exposure must be specified (e.g., 96-h LC50).

Lethal - Causing death by direct action. Death of fish is defined as the cessation of all visible signs of movement or other activity.

Major Habitat Type. A habitat category at or above class level that is widely represented in the site *study* area(s) and is likely to influence the biological response of interest (e.g., habitat features that are major for benthos may not be relevant for fish).

Matrix Spike. A sample prepared by adding a known quantity of a target analyte to a subsample of a real sample in order to determine the effect of the sample matrix on the recovery efficiency of the analyte.

Method Blank. A sample (often distilled water) that does not normally contain the analyte of interest, but is processed with samples containing an analyte of interest through all the steps of the analytical procedure.

Method Detection Limit (MDL). The minimum concentration that must be observed to justify the claim to have detected an analyte with a specified risk (normally 5% of 1%) of making a false detection.

Planktonic. Suspended in the water column of an aquatic habitat and dependent on currents and other water movements.

Precision - the degree of variation in the estimate among individual measurements of the same parameter on laboratory or field replicates using a specific analytical method. The level of replication dictates whether analytical or overall measurement precision is estimated.

Quality Assurance (QA). An integrated system of internal and external activities involving quality planning, quality control, quality assessment, quality reporting and quality improvement to ensure that data meet the needs of users.

Quality Assurance Plan. A document specific to a project or program that outlines its management structure, questions to be resolved, data quality requirements, and QA/QC procedures related to those requirements.

Quality Control (QC). An overall system of internal technical activities whose purpose is to measure and control the quality of the data so that they meet the needs of users.

Quality Control Sample. A sample used to assess the performance of a measurement method in controlling precision and bias.

Quality Management Plan. A document specific to a program or institution that outlines management policies, organizational structure, responsibilities, authorities, and accountability related to QA/QC.

Replicate station - a replicate station is a specific, fixed sampling location within an area, and within which a sample consisting of a minimum of two field sub-samples will be collected.

Sampling area - a sampling area is defined to be a relatively large geographical area which has a defined level of exposure to mill effluent.

Standard operating procedure (SOP) - a written document which details methods of analysis, operation or action whose techniques and procedures are thoroughly prescribed and which are accepted as the method for performing certain routine or repetitive tasks.

Standard reference material (SRM) - a prepared sample of a particular matrix which, when analyzed, is used to determine the accuracy of an analysis.

Sublethal - Detrimental to the organism, but below the level that directly causes death within the test period.

Surrogate Spike - A sample prepared by adding a known quantity of an analyte that mimics the behaviour of the target analyte but is not likely to be found in environmental samples. Recovery of the surrogate is used to estimate recovery efficiencies of the target analytes.

Tainted fish - fish that contains an abnormal odour or flavour.

Toxicity - The inherent potential or capacity of a material to act on a group of selected organisms, under defined conditions. An aquatic toxicity test usually measures the proportion of organisms affected by their exposure to specific concentrations of chemical, effluent, elutriate, leachate, or receiving water.